Utilizing Social Network Analysis in Support of Nation Building

THESIS

Brandon J. Bernardoni, Captain, USAF

AFIT/OR/MS/ENS/11-01

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT/OR/MS/ENS/11-01

Utilizing Social Network Analysis in Support of Nation Building

THESIS

Presented to the Faculty of the
Department of Operational Sciences
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Brandon J. Bernardoni
Captain, USAF

March 2011

AFIT/OR/MS/ENS/11-01

Utilizing Social Network Analysis in Support of Nation Building

Brandon J. Bernardoni

Captain, USAF

Approved:

//Signed//	18 March 2011
Dr. Richard F. Deckro Committee Chairman	date
//Signed//	18 March 2011
Maj. Matthew J.D. Robbins Ph.D. Committee Member	date

AFIT/OR/MS/ENS/11-01

Abstract

Social network analysis is a powerful set of techniques used by social scientists to study the formal and informal interrelations in a community. Since 9/11 these techniques have been increasingly utilized by the defense and intelligence communities to analyze terrorist networks to aid in thwarting foes. This study investigates the use of social networks and structural hole theory to facilitate nation building in failed and failing states. Through the investigation of the underlying social structure of a community, identifying structural holes and gaps within the government or society, Security Stabilization Transition and Reconstruction Operations (SSTRO) efforts can be focused to strengthen the host nation government to provide security and unity for its citizens.

This investigation focused on exploring techniques that link individuals in the professional and governmental community. It was found that Burt's technique of structural holes can be applied to a national level in order to identify structural gaps within an ethnically fractured failing state. This technique can highlight national, regional, or local holes that can be filled to facilitate nation building.

AFIT/OR/MS/ENS/11-01

To the King

Acknowledgements

Praise be to the God and Father of my Lord and Savior Jesus Christ who always leads us in triumphal procession, for the very breath I breath and the life flowing through my veins. He continually gives me strength, and without Him, I am nothing.

Thank you to my beautiful bride and daughter, who tirelessly supported me in this effort. They both listened and aided me in fleshing out ideas, and endured this season of growth right by my side. I love you both infinitely.

I would also like to express my appreciate to my research advisor for his guidance throughout the course of this thesis. He certainly allowed me to learn and grow during this process. I would also like to thank my reader, who gave great feedback to improve this work.

Finally, thank you to the individuals in the Human Terrain System for their help and guidance with my many questions.

What a joy it is to continually learn new things...

Brandon J. Bernardoni

Table of Contents

Page

Page

List of Figures

List of Tables

Table		Page

List of Abbreviations

Utilizing Social Network Analysis in Support of Nation Building

1. Introduction

1.1 Background

The Department of State (DoS) created the Office of the Coordinator for Reconstruction and Stabilization (S/CRS) in 2004. The S/CRS places renewed emphasis throughout the United States Government (USG) on the coordination of capabilities and efforts towards stabilization and reconstruction in post conflict nations [45]. The Department of Defense (DoD) is fully integrated into this effort, especially post 9/11. Current Joint Operations doctrine, or Joint Publication (JP) 3-0 directs Joint Force Commanders (JFC) to,

> "pursue attainment of the national strategic end state as sustained combat operations wane by conducting stability operations independently and/or in coordination with indigenous civil, USG, and multinational organizations" [43:xxiii].

The United States has engaged in conflict and reconstruction operations in failed or failing states for at least the last century. The Institute of Peace and Conflict Studies (IPCS) defines a failed state as,

> "Failed states, alternatively referred to as fragile, weak, quasi or even collapsed states, are seen as weak and ineffectual in providing basic public goods like territorial control, education and healthcare, and legitimate institutions to their populations, and unable, unwilling, or at the worst, complicit in the violation of the fundamental rights of their people . . . their revival is contingent on external intervention or assistance." [40:1].

A current concern of U.S. involvement is to ensure that a country is stabilized once major conflict actions are over and does not recede to a failed or failing status. This achieves two primary objectives, (1) the U.S. and Coalition strategic objectives are accomplished regarding national interests, and (2) the objectives of the third party nation are met, founded on the concept of a viable, sustainable peace. If such a peace can be established and maintained, the state is stabilized and prospers; further international military intervention will not be necessary. Army Field Manual 3-07 states

> "The greatest threat to our national security comes not in the form of terrorism or ambitious powers, but from fragile states either unable or unwilling to provide for the most basic needs of their people... while fostering the development in ways that preclude the requirement for future military intervention" [39:ii]

Covey, Dziedzic, and Hawley offer the following definition for a viable peace,

Definition 1 (Viable Peace): *Peace becomes* viable *when the capacity of domestic institutions to resolve conflict peacefully prevails over the power of obstructionist forces ... viable peace is the decisive turning point in the transformation of conflict from imposed stability to self-sustaining peace [25:xi].*

Viable peace becomes dominant within a nation as insurgent motivations are transformed through the state's ability to recognize, confront, and overcome those motivations. [25]. Current U.S. doctrine outlines the four instruments of power to overcome obstructionist forces, Diplomatic, Information, Military, and Economic (DIME). Gompert, Kelly, Lawson, Parker, and Colloton of RAND argue that sustainable peace cannot be simply accomplished by military means alone, but by a combination of all four instruments of power [32:5]. Often the source of violent conflict is due to an insurgency that has operated within the boarders of the failed or failing nation. Joint Publication 3-24 *Counterinsurgency (COIN) Operations* defines insurgencies in the following manner,

Definition 2 (Insurgency): *Insurgencies are based on internal conflicts focusing on the population. An insurgency attempts to gain power and influence, or promote a particular ideology. The goal of gaining power and influence may not result in overthrowing the host nation government, but by gaining power and influence at a greater rate than other means would peacefully or legally allow [42:II-1].*

Insurgencies thrive in failing and failed states as demonstrated by the spectrum of fragile states outlined in JP 3-24 (see Figure 1.1). The decision for the USG to thwart or lessen this period of vulnerability provides, in part, an incentive for the international community to aid the host nation transition from a failed state to a recovering state. A large part of this transition is conducting operations to counter insurgents within the failed or failing state. According to JP 3-24, COIN is defined as,

Definition 3 (COIN): *COIN is comprehensive civilian and military efforts taken to defeat an insurgency and to address any core grievances. COIN is primarily political and incorporates a wide range of activities, of which security is only one [42:I-2]*

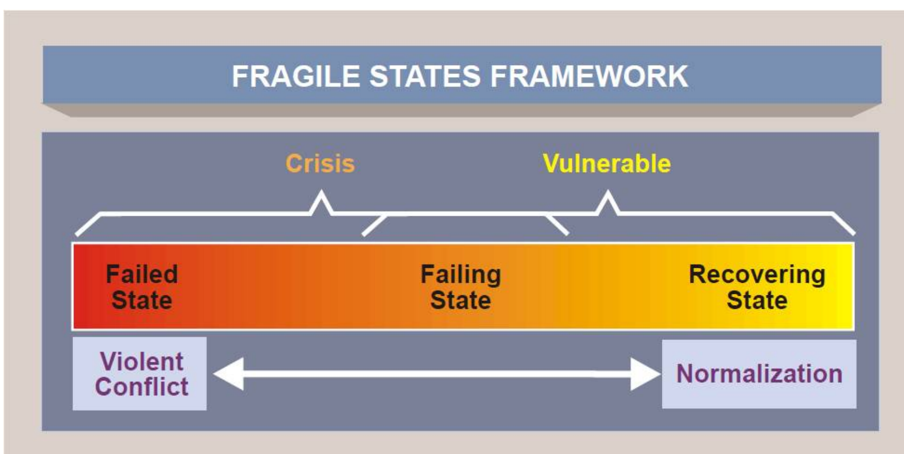


Figure 1.1 Fragile States Framework [42:I-3].

The life cycle of an insurgency can be visualized by Figure 1.2 where the legitimacy of the government is at odds with the strength of the insurgency. The weaker the government, the more likely insurgencies will arise from the population

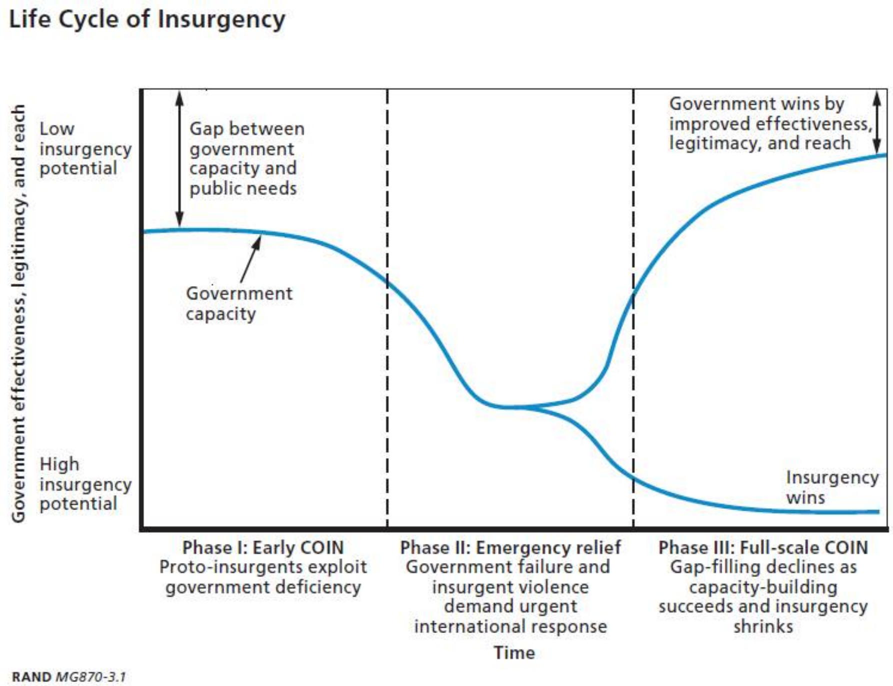


Figure 1.2 Conceptual Insurgency Lifecycle [32:60]

that government serves. Insurgents seeking power and influence capitalize on a weak governments inability to maintain law and order. Therefore, methodologies are needed to examine these issues causing weakened governments.

This thesis examines the stabilization of failed or failing states' government using SNA. A methodology is proposed by which the government may identify citizen demands. SNA is used to identify and direct efforts to close the gap between government capacity and public needs. As Ghani and Lockhart assess, an unseen government is an ineffective government [31]; if citizens do not see action promised by the government then the government is not meeting its basic purpose. The gap indicates an environment of distrust and unreliable security for the people.

1.1.1 Stability, Security, Transition, and Reconstruction Operations. Of primary concern to the U.S. and its Allies is the nation of Afghanistan and it's current state (for example, as a notional measure of the failed condition of Afghanistan, it is currently ranked 176/178 of corrupt countries [5]) leading to it having been, and potentially continuing as, a haven for terrorists. The USG is committed to aiding the Afghan government to develop and sustain a viable peace in the region. Within the DoD nation building and stability operations are formally known as Stability, Security, Transition, and Reconstruction Operations (SSTRO). DoD Instruction 3000.05 defines stability:

Definition 4 (Stability Operations): *A term encompassing various military missions, tasks, and activities conducted outside the U.S. in coordination with other elements of DIME to reestablish or maintain a safe and secure environment, provide essential governmental services, emergency infrastructure reconstruction, and humanitarian relief [29].*

The Chairman of the Joint Chiefs of Staff (CJCS) notes that, "Military support for SSTR are DoD activities that support USG plans for SSTR operations that lead to sustainable peace while advancing U.S. national interests" [26]. Again emphasized in the Joint Operations Concept, "SSTR operations are highly integrated interagency operations that involve a carefully coordinated deployment of military and civilian, public and private, U.S. and international assets" [27:2]. As stated in Definition 1, SSTR operations are vital for the successful implementation of self-sustaining peace in a failed or failing state. It is in U.S. national interest to ensure the viability and sustainability of the government of Afghanistan and other failed and failing states.

In approaching SSTR operations, it is critical to ensure that there is buy-in and cooperation from the national leadership and the support of the people of the host nation receiving international support. Without local popular support, all efforts would be for naught [42:III.1.a].

RAND defines nation building as,

Definition 5 (Nation Building): *"efforts carried out after major combat to underpin a transition to peace and democracy involving deployment of military forces . . . comprehensive efforts to rebuild health, security, economic, political, and other sectors" [44:xv].*

SSTR operations require the precise coordination of instruments of power to assist a failed or failing state. Stabilization activities are to manage tensions that deteriorate security, economic or political systems and establish preconditions for reconstruction efforts. Security establishes safe environment for populace, government and agencies conducting SSTR operations. Transition is the process of shifting lead responsibility and authority from intervening agencies to host nation. Finally, reconstruction is the process of rebuilding political, socioeconomic, and infrastructure for long term development [27:2].

1.1.2 Social Network Analysis. Social Network Analysis (SNA) has been developed by social scientists over the years to study the social structures of relationships. Methods, theory, empirical research, and formal mathematics and statistics have developed to understand the importance of relationships amongst people [65:3]. The internet generation has popularized the term *social network* with the use of Facebook©, Twitter©, MySpace©, and many other networking websites. SNA is *not* social networking. Social networks provide tools for descriptive analysis in order to understand the structure of a network. SNA applies mathematical, computer science, and operations research tools and methods, in addition to social science theory, for prescriptive analysis. Such prescriptive analysis attempts to measure and represent these structural relations accurately and to address not only why they occur, but their subsequent projected consequences [47]. SNA provides a rigorous method of analyzing the interactions between members of a group. Utilizing graph theory concepts, the power and speed of network analysis can be quickly leveraged to analyze relationship structures between individuals.

SNA has broad application and has been used to help streamline business processes, improve coordination across agencies, identify terrorist networks, as well as quantify a large portion of the internet [65, 19, 15, 60]. In the context of organizations and business collaboration, SNA can help identify the informal structure of the organization and influence growth by developing these informal networks [64:109]. The implications of SNA for business managers include a better understanding of informal organization, the capture and exploitation of new ideas, efficient dissemination of information, and to more effectively understand working habits of employees [12:179].

The use of SNA has not yet directly expanded to the issue of nation building through application to a governmental structure. This thesis investigates such applications.

1.2 Problem Statement

There is a need to have methodologies in place at the tactical, operational, and strategic level of nation building to determine a viable economic, social, and governmental sustainability while recognizing the cultural differences of the U.S., Coalition and the third party nation. In this thesis, SNA is used to identify gaps between government capacity and public needs. These gaps are referred to as *structural holes* in the SNA literature [19]. Identifying these structural holes will bring to light areas in which to focus resource investment for the re-building of a post conflict nation. This approach is applied to a notional open source model of the Afghan national government as a demonstration of the technique.

Gompert *et al.*, in a RAND study, defines four areas of civilian COIN [32:61]:

- *Indigenous capacity building:* e.g., public sector reform, infrastructure refurbishment, training and public sector development of doctors and teachers.
- *Public-service gap-filling:* e.g., public education, population security functions, public heath services, justice and civil administration.
- *Fostering development to create livelihood opportunities:* e.g., job training and placement, direct investment and marketplaces, production and distribution links.
- *Emergency humanitarian relief:* e.g., supplying dire needs with water, food, shelter, sanitation and urgent medical care.

This thesis focuses primarily on extensions of "public service gap filling" described by Gompert *et al.* In order to understand how to fill gaps, one must first identify where those gaps exist. This thesis utilizes methodology to identify gaps within the government capacity to aid in building the legitimacy of the host government. The outlined approach, however, could be applied to any social structure in a nation building effort.

One aspect of "public-service gap-filling" is that of the professional corps (i.e., the lawyer, doctor, businessman, even religious leaders). One cannot begin to build a sustainable government if professionals cannot live as productive members of their given society. As Greg Mortenson, the director of the Central Asia Institute (CAI), an organization building schools throughout Pakistan and Afghanistan, states

> "once you educate boys, they tend to leave the villages ... but the girls stay home, become leaders in the community, and pass on what they've learned. If you really want to change a culture, to empower women, improve basic hygiene and health care, and fight high rates of infant mortality, the answer is to educate girls" [49:209].

He also points out that "one must water a plant before it could be coaxed to grow; children had to survive long enough to benefit from school" [49:201].

There is a need to combine military operations with political and economic development as part of a single comprehensive and coordinated campaign to convince the people of a nation that the legitimate government is a better option than an insurgency [32:5]. Trust in a sustainable, predictable national government brings about security and stability as individuals in a local area know that the government is more interested in their greater good than that of the local terrorist or insurgent regime. The challenge then is how people of many different relations, backgrounds, and ethnicity unite on a common front to achieve the greater goal of building a nation?

For years, the business community has been trying to answer similar questions although on a different scale and focus. Instead of trying to build a nation to provide stability and security in exchange for power, business is trying to build an organization to provide goods or services in exchange for increased stakeholder wealth. A nation is defined as "a politically organized nationality (or body of people of several different nationalities) or politically organized society having particular character and the operations therein" [66:773]; a business is defined as "commercial or mercantile activity engaged in as a means of livelihood; concerned with the purchase or sale, supply and distribution of commodities or in related financial transactions" [66:154]. It is a premise of this study that what businesses have done and learned in the past to build their organization utilizing the inherent social structure can be leveraged, translated and applied to the national level in order to help build a legitimate national government.

1.3 Methodology and Approach

The proposed methodology extends efforts previously found in business and economic SNA models and applies them to building a national level government. Data was collected via open source documents over the internet and through open-source databases collected for intelligence personnel such as the Complex Operations Wiki sponsored by the Human, Social, Cultural and Behavior Modeling (HSCB) and the DoD [4].

1.4 Research Scope

This thesis focuses primarily on what JP 3-0 defines as Phase IV (see Figure 1.3), to stabilize and establish security and restore services and Phase V, to enable civil authority and redeployment [43:IV-29]. The lines are blurred between phases IV and V; in order to stabilize, the military has to ensure that the threat is reduced to a manageable level such that the oncoming civil authority can lead the country. However, it can also be argued that in order to have security, the nation must be stable with civil authority in place. For example, Afghanistan is still in the midst of securing the environment such that the civil authority in place can mitigate the effects of the current threat level. The level of security may also vary in different areas of a nation.

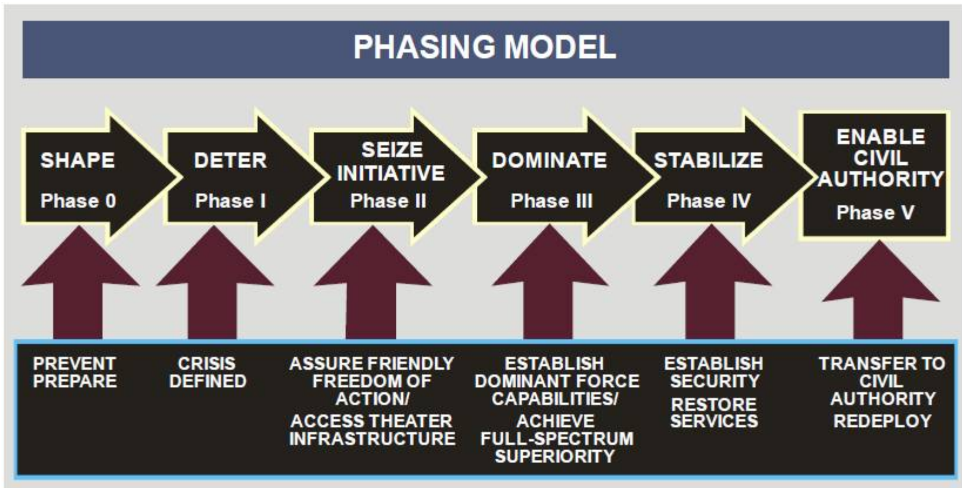


Figure 1.3 JP 3-0 Phasing Model [43:IV-27]

This thesis focuses on the "fostering economic stability and development" [29:4.d.4] aspect of SSTRO. Specifically the identification of gaps between government legitimacy and public needs. Legitimacy includes aspects of corruption, the ability to provide basic services such as health, water, food, and shelter, the ability to foster economic development, and the inclusion of all members of the nation in the civic process. Needs are defined by the people themselves as outlined in the Constitution of the Islamic Republic of Afghanistan [1].

1.5 Assumptions

It is important to address several key assumptions. First, SNA is applicable to government and non governmental structures alike. Second, Structural holes exist and are identifiable with current SNA techniques. Third, structural relations bring more insight to explain behavior of individuals when compared to individual attributes (e.g., age, gender, race) [47:4]. Fourth, social networks and social structure have impact on the beliefs and actions of individuals within that network [47:5].

Fifth, while structural relations should ultimately be viewed dynamically, this thesis looks at a snapshot of past and current data as a static relationship. Sixth, a nation-state functions in a similar fashion to a large business conglomerate, where multiple corporations engaged in different businesses are under one corporate structure [34: "conglomerate"]. Finally, a nation's advancement and profit (i.e., ability to effectively utilize human resources) is tied directly to the concept of social capital.

1.6 Overview

The thesis is organized as follows: Chapter 2 provides a review of the business and economic literature in which SNA techniques are applied to the construction of organizations. Moreover, a review of relevant SNA concepts is provided. Chapter 3 presents a description of how data is collected and organized for this study. Additionally, a test case is examined demonstrating the methodology presented. Chapter 4 describes how countries are listed as failed states. Then background of the past and current condition of Afghanistan is presented. Finally, analysis is conducted on a demonstration data set for the government of Afghanistan. Chapter 5 concludes the thesis and provides findings and applications to SSTRO.

2. Literature Review

2.1 Introduction

This chapter presents a brief literature review to include concepts key to SNA and a summary of SNA literature within the business community that was applied to building a post-conflict nation. Then a review of failed and failing states and how to stabilize them through COIN operations followed by a review of current literature containing different views of how to build a nation will be presented.

2.2 Social Network Analysis

2.2.1 A Brief History of SNA. SNA is a scientific area focused on the study of relations, often defined as social networks [65]. In its basic form, a social network is a network where the nodes are people and the relations (also called links or ties) are connections. These connections may be due to friendship, business relationship, kinship and so forth. SNA uses graph theory concepts and applies them to representing the social world [65:7]. SNA is also called network analysis, structural analysis, the study of human relations, and is often referred to as the science of connecting the dots [21]. Ultimately, SNA draws from the fields of Psychology and Sociology to study people and the relationships between groups of people [65:8].

Today, SNA refers to the analysis of any network in which all of the nodes are of one type (e.g., all people, or all roles, or all organizations), or multiple types (e.g., people and the groups they belong to) in order to understand the society itself. The metrics and tools in this area, since they are based on the mathematics of graph theory, are often applicable regardless of the type of nodes in the network or the reason for the connections, given the context of the problem [21].

SNA is rooted in the work of J.L. Moreno's sociogram [65:77]. This concept evolved into the analytical approach of sociometry, later becoming a very useful notation for social networks. The sociogram represents people as points in twodimensional space with relationships as lines between these points. SNA has developed out of multi-disciplinary efforts and has a broad appeal due to the social network perspective [65:77]. This interdisciplinary behavioral science is grounded in the observation that social actors are interdependent and the links between them have important consequences for all involved.

Wasserman and Faust describe SNA as a

> "distinct research perspective within the social and behavioral sciences; distinct because SNA is based on an assumption of the importance of relationships among interacting units" [65:4].

Typically, social analysis focuses on individual attributes. SNA focuses on ties, relationships, and interactions to gain insight on behavior between individuals. Statistical and descriptive techniques are used within SNA, however different assumptions are required for statistical analysis [65:5,8].

For most researchers, the nodes are actors. As such, a network can be a cell of terrorists, employees of global company, or simply a group of friends. However, nodes are not limited to actors. Nodes can be individuals, groups, large organization, or even nation-states [47:4]. A series of computers that interact with each other or a group of interconnected libraries can also comprise a network, as can how various resources are interlinked with one another [21].

2.2.2 Definitions of SNA. A social network is a social structure made of up individuals connected to other individuals by friendship, kinship, business transaction, and so forth. SNA is *not* social networking on Facebook©. On the contrary, SNA views social relationships in terms of network theory. SNA explicitly assumes that actors participate in social systems connecting them to other actors, causing influence upon one another leading to a desired set of actions [65]. The methodology of SNA attempts to measure and test hypotheses about how these relations influence other actors within the network. According to Wasserman and Faust,

> "SNA provides a precise way to define important social concepts, a theoretical alternative to the assumption of independent social actors, and a framework for testing theories about structured social relationships" [65:17].

Several key definitions within the SNA literature, taken primarily from Wasserman and Faust [65], follow to convey basic SNA concepts. Beginning with the most basic unit of a social network graph, the actor is defined as:

Definition 6 (Actor): *"discrete individual, corporate, or collective social units." [65:17]*

Actors can be people in a group, departments in a corporation, public service agencies in a city, or even nation states [65:17]. Actors are objects within a graph and are also known as nodes in graph theory literature.

How actors relate with one another, or their interaction is a *relational tie*.

Definition 7 (Relational Tie): *"Actors are linked to one another by social ties. The defining feature of a tie is that it establishes a linkage between a pair of actors" [65:18].*

Some examples of ties are friendship, liking, respect, business relations (transfer of resources), association, movement between locations, physical connection (like a bridge or road), formal relations (hierarchical structure), or familial relations [65:18].

There are several types of ties between two or more actors. These are broken out into *dyads*, *triads*, *subgroups*, and *groups*.

Definition 8 (Dyad): *"A linkage or relationship establishing a tie between two actors. The tie is inherently a property of the pair and therefore is not thought of as pertaining simply to an individual actor. The basic unit for the statistical analysis of social networks" [65:18].*

Definition 9 (Triad): *"A subset of three actors and the possible ties among them" [65:19].*

Definition 10 (Subgroup): *"Any subset of actors, and all ties among them" [65:19].*

SNA is not just a matter of putting people into dyads, triads, subgroups, and groups, but more importantly it is being able to model relationships within an entire system. Thus Wasserman and Faust define a group as,

Definition 11 (Group): *"The collection of all actors on which ties are to be measured . . . arguable by theoretical, empirical, or conceptual criteria that the actors actually belong together" [65:19].*

With groups, SNA now becomes an issue of being able to clearly define the boundaries of the network and what, exactly, defines the group to analyze.

Definition 12 (Relation): *"The collection of ties of a specific kind among members of a group . . . a relation refers to the collection of ties of a given kind measured on pairs of actors from a specified actor set. Ties themselves only exist between specific pairs of actors" [65:20].*

For example, friendships or diplomatic relations between nations might define relations. Analysts may measure several different relations for a given group of actors. Relations can be either *directional* with one actor relating to another or *non-directional* with mutual interaction within the relation.

There is also significance when relations exist between groups.

Definition 13 (Bridge): *"A line [edge] that is critical to the connectedness of the graph" [65:114].*

A bridge is an edge that, if deleted, would cause the endpoints of the edge to be wholly contained within different groups of the graph. More in terms of SNA, Burt defines a bridge as,

Definition 14 (Bridge (Burt)): *"The relation between two [actors] is a bridge if there are no indirect connections between the two people through mutual contacts" [20:294].*

The complement of a bridge is a cohesive subgroup. Graph theory provides definitions for cohesive subgroups in social networks.

Definition 15 (Clique): *"A nondirectional dichotomous relation, a maximal complete subgraph of three or more nodes. It consists of a subset of nodes, all of which are adjacent to each other and there are no other nodes that are also adjacent to all the members of the clique" [65:254].*

Finally, a social network can now be formally defined,

Definition 16 (Social Network): *"Consists of a finite set or sets of actors and the relation or relations defined on them. Relational information is critical to defining of a social network" [65:20].*

Knoke offers a slightly different definition of social network,

Definition 17 (Social Network (Knoke)): *"A structure composed of a set of actors, some of whose members are connected by a set of one or more relations" [47]*

J. Clyde Mitchell provides a third definition,

Definition 18 (Social Network (Mitchell)): *"A specific set of linkages among a defined set of persons, with the additional property that the characteristics of these linkages as a whole may be used to interpret the social behavior of the persons involved" [48]*

SNA takes into account both the presence and absence of ties among actors, and possibly even variations in strength of those ties. It also proposes that because networks can be viewed at the individual and system level, network analysis can explain changes in structural relations and their effects [47:9]. One would model these relationships in order to obtain a picture of how the group is structured, then study how the structure of the group impacts the function of the group or how the structure influences other individuals within the group [65:9].

Due to the importance of relations between actors within SNA, Knoke offers a taxonomy of relational forms that take place within networks, listed more or less in terms of increasing relational intensity [47:12]:

- *Transaction* relations, as in business sales and purchases.
- *Communication* relations, links through which messages can be sent, e.g., an executive officer.
- *Boundary penetration* relations, members overlapping in two or more social structures, e.g., a member who is in research and development and is IT support.
- *Instrumental* relations, actors contact one another in efforts to secure valuable goods, services, or information, e.g., the "good old boy system".
- *Sentiment* relations, actors expressing their feelings of affection, admiration, difference, loathing or hostility toward one another.
- *Authority/power* relations, indicate the rights and obligations of actors to issue and obey commands, e.g., formal hierarchical organizational structure.
- *Kinship and descent* relations, bonds of blood and marriage.

Differing ties between actors in a network create a general feeling of people to strive for social balance and harmony within their relationships thus making them more apt to act within the groups established norms. This is the foundation of building trust, cohesion, and cooperation while simultaneously reducing individualism and power concentrations [22:42,461-464]. Network structure is important to trust, influence, reciprocity, and leadership, thus SNA is a powerful tool suitable to study social interactions [60:109].

2.2.3 Network Types and Graph Theory. There are several different types of social networks that can be observed and studied. Wasserman and Faust categorize networks by the sets of actors and the associated ties among them. Mode of a network is defined as,

Definition 19 (Mode): *"The number of sets of entities on which [ties of a specific kind between pairs of actors] are measured" [65:35].*

One-mode networks are most common and focus on a single set of actors. Actors can be various entities such as people, subgroups, organizations, or even nation-states, progressing from the individual to collections of individuals, and to aggregates of subgroups of people [65:36]. Relations can be defined as Knoke describes in his topology listed previously.

Two-mode networks focus on two sets of actors, or even one set of actors and one set of events. These *dyadic* networks are functions of dyads of which the actors are from different sets or subgroups [65:39]. Additionally, an *affiliation network* is where there is one set of actors and one set of events where actors are measured by their association to an event or activity. There are higher-level mode networks but are not often used in SNA [65:35].

Another design of network is an *ego-centered* network. This type of network focuses on a single actor, called *ego*, and a set of *alters* who are relationally tied to ego [65:42].

These definitions are rooted in social science and graph theory, but there are several statistical and social assumptions that underly the mathematics behind SNA.

2.2.4 Assumptions of SNA. Knoke lists three primary assumptions within SNA. The first assumption for SNA is that the structural relations between actors are often more important in understanding behaviors of those individuals than attributes such as age, race, gender, ideology, sex, and so forth. When the analysis focused solely upon actor attributes many important explanatory insights provided by network perspectives on social behavior can be lost. Knoke also mentions that many attributes remain unaltered in different social context whereas structural relations *can* change depending on context. The second assumption for SNA is that social networks affect perceptions, beliefs, and actions. In other words, the more time spent with a person, the more influence each actor will have on the other's belief structure. According to Knoke, these structural relations are a double edged sword; crucial to sustaining cohesion and solidarity within a group, but this can also reinforce prejudices. Finally, SNA should be viewed as a dynamic process, acknowledging that networks are not static and develop over time, although most analysis looks at a single snapshot in time [47:4-6].

Wasserman and Faust note that SNA methods generally assume finite actor set size and also an enumerable set of actors. These conditions emphasize the importance of boundaries on the data set. It is also assumed that relevant information on all important actors can be collected [65:32-33].

A number of important principles distinguish network analysis from all other related social science. According to Wasserman and Faust [65:4]:

- Actors and their actions are viewed as interdependent rather than independent, autonomous units.
- Relational ties (linkages) between actors are channels for transfer or flow of resources (either material or nonmaterial).
- Network models focusing on individuals view the network structural environment as providing opportunities for or constraints on individual action.
- Network models conceptualize structure (social economic, political and so forth) as lasting patterns of relations among actors.

2.2.4.1 Statistical Assumptions. Wasserman and Faust note that statistical methods of analysis differ from structural analysis in that there is knowledge of underlying models based on probability distributions [65:505]. Van der Hulst argues that inferential statistics such as hypothesis tests and significance tests become problematic in SNA because the basic assumptions of independence, non-random sampling, and unknown distribution of variables are not met [60:110]. However, some of these problems can be overcome by the use of permutation tests [60:110]. Van der Hulst also could suggests one could employ the use of non-parametric statistics. Parametric statistical methods apply to problems where the distribution of samples taken is known; non-parametric applies to unknown sample distributions, as is the case in SNA [63:742].

2.2.5 SNA Data. Wasserman and Faust describe SNA network data as,

> "data consists of at least one structural variable measured on a set of actors. The substantive concerns and theories motivating a specific network study usually determine which variables to measure, and often what techniques are most appropriate for their measurement" [65:28].

There are two types of variables that contain data, *structural* and *composition* [65:29]. Structural variables do exactly what their name implies, gives structure to the network by measuring ties of a specific kind between pairs of actors, also seen as interactions of varying degrees. Composition variables measure actor attributes on the individual actors themselves, such as ethnicity, race, gender or location [65:26].

There are several ways to collect SNA data including conducting interviews and surveys, observation of actors within the boundary set, or data extraction from archival records [47:21-31]. This thesis focuses on data extraction from archival records; this approach is explained in more detail in § 3.3.2. Data in SNA focuses primarily at the unit level.

The *modeling unit* is the level (actor, dyad, and so forth) at which network data is studied. This data is either directional (*dichotomous*) or nondirectional (*valued*) [65:43]. For example, a study on the importance of individuals in an organization may count the number of emails each individual receives and writes in the course of a day, while recording the attributes such as age, gender, and frequency of emails from a particular individual. Hypotheses of relations are made based on the collected data and the importance of that data.

As a second example, consider a SNA approach examining who is calling whom. In particular, the SNA approach examines the importance of the individual versus who is making the most calls. This is an effort to differentiate between who is the secretary and who is the commander. The secretary makes many more phone calls than the commander, but it is the commander who is in charge of the unit, not the secretary. SNA, when considering more than one structure, provides tools to understand who is important within a network. The social network perspective looks at the relationships between the actors and projects network structure based upon actor interactions.

Social network data is often displayed in a sociomatrix (see Table 2.1). In the sociomatrix, also known as an *adjacency matrix*, **A**, the presence of a 1 in the matrix indicates a tie or a relation between actor i and j and a 0 represents no tie between these two actors. Let **A** be the adjacency matrix, then,

$$a_{ij} = \begin{cases} 1 & i^{th} \text{ node is adjacent to the } j^{th} \text{ node} \\ 0 & \text{otherwise} \end{cases}$$

From Table 2.1, notice how there is a tie from 1 to 2, but the tie is not reciprocated from 2 to 1. This indicates directional ties where actor 1 may call actor 2 for advice but 2 would not call 1. For undirected ties, the sociomatrix would be a symmetric matrix with reciprocated ties (i.e., $a_{ij} = a_{ji}, \ \forall i, j$). Data can also be represented as strength of tie as well (see Table 2.2). For example, actor 3 has a strength tie with actor 4 of 4, but actor 4 has a tie with actor 3 of 2. This relational matrix is easy to conceive but may be difficult to develop.

An alternative way for social network data to be displayed is a node adjacency list, listing all nodes that the initiating node interacts with, as seen in Table 2.3. Beginning with the left most node, there is a direct tie between all nodes following on that line. For example, actor A is relationally tied to actors B, E, F and U.

The node adjacency list can be simplified one step further into an edge list. The edge list contains a separate listing for every edge that a node contains, representing only one edge at a time. For example, the first line in Table 2.3 would have an edge list of $\{(A, B), (A, E), (A, F), (A, U)\}$, and each subsequent row afterwards.

Table 2.1 Example of a Sociomatrix [65:740]

Manager	1	2	3	4	5	6	7	8	9	10
1	0	1	0	1	0	0	0	1	0	0
2	0	0	0	0	0	1	1	0	0	0
3	1	1	0	1	0	1	1	1	1	1
4	1	1	0	0	0	1	0	1	0	1
5	1	1	0	0	0	1	1	1	0	1
6	0	0	0	0	0	0	0	0	0	0
7	0	1	0	0	0	1	0	0	0	0
8	0	1	0	1	0	1	1	0	0	1
9	1	1	0	0	0	1	1	1	0	1
10	1	1	1	1	1	0	0	1	0	0

These methods of importing data can also be viewed as graphical representations. It is possible to construct a visualization of the actors' relationships with one another, with each actor represented by a node and each relationship represented by an edge. Using NetworkX [35], Table 2.3 is represented graphically in Figure 2.1.

2.2.5.1 Difficulty in Gathering SNA Data. Information provided by SNA is very useful to organizations and individuals alike, however collecting this data is far from easy [65]. This data collection is complicated when gathering data

Table 2.2 Example of Directed, Weighted Sociomatrix [65:745]

Manager	1	2	3	4	5	6	7	8	9	10
1	0	4	2	2	2	2	2	2	2	2
2	4	0	2	0	1	0	3	3	4	1
3	3	1	0	4	1	0	0	2	0	2
4	2	0	2	0	2	0	0	2	2	2
5	3	0	0	2	0	0	0	2	3	2
6	3	0	0	0	0	0	0	2	0	0
7	3	2	1	0	0	0	0	2	2	0
8	2	2	2	2	2	0	0	0	1	0
9	3	4	0	0	2	0	0	2	0	0
10	2	1	3	3	2	0	1	2	2	0

Table 2.3 Node Adjacency List [19:56]

Adjacent Nodes						
A	B	E	F	U		
B	A	D	U			
C	U					
D	B	U				
E	A	U				
F	A	U				
U	A	B	C	D	E	F

of large populations (i.e., a national level) over large periods of time (i.e., several decades).

Of the various methods of data collection described in §2.2.5, the most resource intensive is that of a personal interview where individuals interact on a regular basis [65:48]. In addition, it requires a working knowledge of who is being interviewed. This method may not be practical for analyzing data on a national level.

The resources that are required to observe actors within a nation can be prohibitively costly and time consuming. Observation, questioning, and interviewing actors requires permission; and that permission is not always obtained [65:43,49].

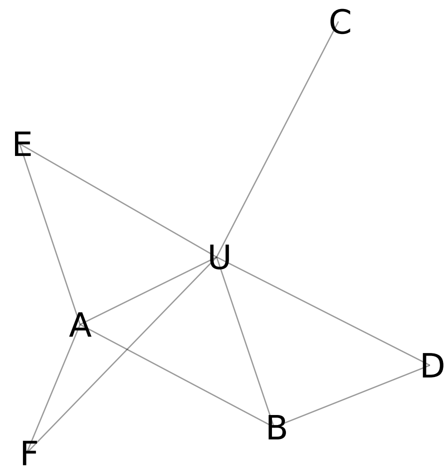


Figure 2.1 Graphical Form of Table 2.3 [19:56]

Finally, extracting data from archived records such as news articles, published documents, internet sources, court records, minutes of executive meetings, publicly available information on the nation, and so forth [65:50-51], are all possible methods for data collection on a failed or failing state. For example, Burt obtained information on interactions among corporate actors and executives from the 1950's on the front pages of old *The New York Times* newspapers and reconstructed ties from the past [19]. Today, there is vast amounts of information posted on the internet. An analyst can continue to build their data set over a number of years, however most research programs and studies are limited by time.

All of these collection methods are an attempt to postulate ties between actors and are common amongst social and behavioral science methods [65:51].

2.2.6 SNA Measures. This section presents a description of different measures from SNA that are used within this thesis. This development focuses primarily on who will be important in a network in order to identify the structural holes within the network. While these measures are at the individual actor level, they can be aggregated to higher level networks. Wasserman and Faust note that the definitions of importance and prominence of actors in a network are the focus of several measures [65:169].

Network measures often require the understanding of different network dimensions including size, density and hierarchy.

Definition 20 (Size): *"Network size, $\mathscr{M}$, is the number of contacts in a network" [20:296]*

A second measure of networks is density.

Definition 21 (Density): *"Density is the average strength of connection between contacts, $\sum a_{ij}/[N*(N-1)]$, summing across all contacts i and j [20:296]*

Density suggests how tightly condensed the connections within a network, $\mathscr{N}$. The third measure of network dimension is hierarchy,

Definition 22 (Hierarchy): *"Another form of closure in which a minority of contacts, typically one or two, stand above the others for being more the source of closure" [20:296].*

Often times, hierarchical networks are built around the supervisor or director. These three network dimensions play a role in the measures to be defined in the following sections.

2.2.6.1 Centrality. Degree Centrality focuses on identifying the actor location within the social network. Wasserman and Faust describe centrality as "prominent actors are those that are extensively involved in relationships with other actors, with the importance that the actor is simply involved" [65:173]. Hanneman adds that, "actors who have more ties to other actors may be in advantaged positions" [37:ch.10]. Simply put, the *actor degree centrality* measures the extent to which an actor is connected to other actors in a social network and addresses the question "Who is involved in many ties in this network, and thus most visible to others?" based on a personal direct connection. In a non-directed graph with g actors, degree centrality for actor i is the sum of i's direct ties to the $g-1$ other actors, also seen in Eq. (2.1).

$$C_D(N_i) = \frac{\sum_{j=1}^{g} x_{ij}}{g-1}, \quad i \neq j \tag{2.1}$$

Here, $C_D(N_i)$ denotes degree centrality for node i and $\sum_{j=1}^{g} x_{ij}$ counts the number of direct ties that node i has to the $g-1$ other j nodes and normalizes it by dividing by maximum possible degree in graph $\mathscr{G}$. Simply put, this is adding all the cell entries in either the given actors row *or* column of an adjacency matrix and dividing by the number of alters. It is calculated over each node.

Wasserman and Faust go on to define the general centralization index, or *group degree centralization* presented in Eq. (2.2) to be normalized between 0 and 1. It indicates if one actor has more prominence in a group compared to others within the group.

$$C_A = \frac{\sum_{i=1}^{g} [C_A(n^*) - C_A(n_i)]}{max \sum_{i=1}^{g} [C_A(n^*) - C_A(n_i)]} \tag{2.2}$$

Group degree centralization is also a rough measure of how unequal the individual actor values are including variability, dispersion, or spread of prominence by comparing between actors in the network, $\mathscr{G}$. This can be compared to measures of dispersion in descriptive statistics, such as the common standard deviation [47:64].

Another measure is the *betweenness centrality*. This measure indicates how actors control the relations between unconnected dyads [47:67]. This measure indicates who has control over resource and information flow within a network. It is given by Eq. (2.3). Vertices that occur on many shortest paths between other vertices have higher betweenness than those that do not. Betweenness centrality assumes that graphs are undirected and connected allowing for loops and multiple edges.

$$C_B(N_i) = \sum_{j<k} \frac{g_{jk}(N_i)}{g_{jk}} \tag{2.3}$$

Here, g_{jk} is the number of geodesic (or shortest in terms of number of edges traveled) paths between nodes j and k while $g_{jk}(N_i)$ is the number of geodesics between j and k that contain node i. This simply measures the extent to which node i is along the geodesic path of other members in the network [47:68]. An index of 0 indicates i is on no geodesic path for all pairs in network. The maximum value possible is the total number of actors in the network. Eq. (2.3) can be standardized between 0 and 1 by Eq. (2.4):

$$C'_B(N_i) = \frac{C_B(N_i) \times 2}{(g-1)(g-2)} \tag{2.4}$$

2.2.6.2 Eigenvector Centrality. Bonacich proposed *Eigenvector centrality* as a superior measure of the *relative* importance of a node within a given network $\mathscr{G}$. In other words, the eigenvector centrality is a more sophisticated version of the centrality concept discussed in § 2.2.6.1. Degree centrality is a sum of all connections a node may have, while eigenvector centrality accounts for the fact that not all connections are considered equal [16:555]. Bonacich argues that being connected to highly connected alters makes an actor central in the network. For example, a general officer in the Air Force has more influence than an airman in any given network $\mathscr{G}$ of DoD members, because it is generally assumed he knows more "powerful and influential" individuals than the airman does.

$$C_\lambda(N_i) = \frac{1}{\lambda} \sum_{j=1}^{n} A_{ij} x_j \tag{2.5}$$

In Eq. (2.5), $C_\lambda(N_i)$ represents the centrality of node i, and A_{ij} represents the adjacency matrix of the network. λ is an eigenvalue for which an eigenvector solution exists. For the i^{th} node, the centrality score is proportional to the sum of the scores of all nodes connected to it [16:556]. This can also be represented in matrix or eigenvector notation as $\lambda x = Ax$. This measure is used in Google© page rank [13].

Bonacich suggests that eigenvector measures of centrality have several advantages over conventional graph-theory based measures of centrality. They can allow for variations in the degree to which status is transmitted from position to position and they also account for more than just the degree of a node [16:563].

2.3 Structural Holes

This section contains excerpts and adaptations from Burt's 1992 book, *Structural Holes: The Social Structure of Competition* [19]. Structural holes are a social capital concept suggested by Burt, which is "a broad term encompassing the norms and networks facilitating collective action for mutual benefit" [68:155]. Social capital is more or less based on the fact that individuals or groups gain some advantages simply by their location in the social structure. Burt defines social capital as,

Definition 23 (Social Capital): *"the manner in which resources available to any one person in a population are contingent on the resources available to individuals socially proximate to the person" [19:12].*

In other words, people who do better are somehow better connected; being in certain positions in a network is an asset in its own right. There is advantage in structure of an actor's network, $\mathscr{G}$, and the location of the actor's contacts within $\mathscr{G}$. Actors in well structured networks received higher rates of "return", in that they are more easily connected to ideas, resources, and information. Burt builds on this to say that there is some difference between non-redundant contacts, or two individual contacts from two different social structures [19:18-20,50-53] and there are strategic benefits associated with being in a unique bridging position between other actors. Formally, Burt defines redundancy as,

Definition 24 (Redundancy): *"Players i and j are structurally equivalent* [or redundant] *to the extent that they have identical relations with every other player q" [19:50].*

Mathematically, redundancy is defined at Eq. (2.6a)

$$p_{iq} m_{jq} \tag{2.6a}$$

Where p_{iq} is the proportion of i's network time and energy invested in the relationship with q, represented by Eq. (2.6b) with A_{ij} as the adjacency matrix of network ties from i to j.

$$p_{iq} = \frac{(A_{iq} + A_{qi})}{\sum_j (A_{ij} + A_{ji})}, \quad i \neq j \tag{2.6b}$$

m_{jq} is the marginal strength of contact j's relation with contact q. m_{jq} is mathematically seen as Eq. (2.6c) where $\max(A_{jk})$ is the largest, or strongest, of j's relations with any other member in the network, $\mathscr{G}$. This applies only to directed and weighted adjacency matrices and is 1 in undirected, non weighted matrices.

$$m_{jq} = \frac{(A_{jq} + A_{qj})}{\max_k (A_{jk} + A_{kj})}, \quad j \neq k \tag{2.6c}$$

By way of counter example, two individuals in a network are structurally equivalent if they share the same relations to the same people, or are able to get the same information from multiple sources. Thus, Burt defines a structural hole as,

Definition 25 (Structural Hole): *The separation between non-redundant contacts. Non-redundant contacts are connected by a structural hole [19:18-19].*

Structural holes act as a void between two contacts, and when filled the two contacts add additional value to the network versus reproducing that which is already present [19:18]. For example, cliques have zero structural holes as, by definition, everyone is connected, there is no impedance to the flow of information from one individual to another. Burt argues that individuals who are located in, or who fill, structural holes in networks enjoy certain advantages of information flow in two or more networks, increasing the potential for creativity and new ideas from one network to another. This methodology emphasizes the importance of open rather than closed networks. Burt argues that the networks with the highest economic return lie between and not within regions of relationships. This between networks concept provides opportunities for great economic payoff between different firms, or different departments within the government. Burt says that partner selection, more than social capital, determines effective cooperations between firms [19:16]. A pictorial representation of the structural holes argument can be found at Figure 2.2.

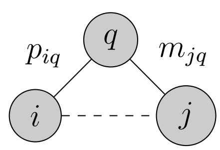
Burt continues his argument that structural holes are opportunities to broker connections between people. There are three terms he delineates:

> "Access to structural holes is discussed as synonymous with brokerage opportunities, both of which are discussed as synonymous with brokerage. All three terns are about the advantage created when connections are made between disconnected people, connections in terms of coordination between the disconnected people, or connections in terms of ideas or resources derived from exposure to contacts who differ in opinion or the way they behave" [20:293-294].

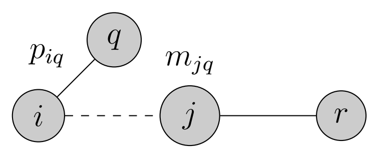
Building on Burt's work, Frankort [30] argues that a firm achieves higher innovative performance by maintaining alliances to others that are not directly connected with one another. The lack of direct connections amongst partners bring about structural holes meaning that a particular firm being aware of, and filling these holes, is more likely to develop more innovative ideas and access to resources than their competitors [30]. This concept is applied to building failed and failing nation states in this thesis, based on the aggregate assumption that a nation-state functions organizationally in a similar fashion to a large multinational corporation. Within these states, there are usually many tribes, factions or power blocks. Where loose connections exist one can observe these structural holes in order to develop programs or procedures to aid in uniting the various factions with the nation during post major conflict reconstruction operations.

Berardo used the concept of structural holes to consider the question "Do organizations with more collaborative partners perform better than organizations with fewer partners?" [15:521]. Berardo proposed a technique to measure how the actors in a project can span structural holes in the network of inter-organizational collaboration. Berardo uses the measure efficiency, defined in Eq. (2.8).

Figure 2.2 gives Berardo's example of structural holes. In Figure 2.2a, actor i establishes a link to j. This new relationship does not span a structural hole since q already has a pre-existing relationship with both i and j. In Figure 2.2b, the new relationship does span a structural hole for q to r.



(a) i does not span structural hole



(b) i spans structural hole

Figure 2.2 An Illustration of the Structural Holes Argument [15:529]

Van der Hulst notes how actors who span structural holes create competitive benefits for themselves and for the network through both information benefits and control benefits between parties [60:108]. Information benefits include early access to non-redundant information which creates opportunities to react accordingly. There are control benefits because the actor who fills structural holes controls information flow between groups. Therefore, actors who fill structural holes tend to be more creative, realize opportunities, and know where to find the right individuals [60:108].

Structural holes capitalizes on the strength of weak ties argument [33:1363-5]. The weak ties argument was proposed by Granovetter in his research linking current network structures to job searches. He hoped to find that people were able to find jobs through close contacts, however he found that people rarely found work through these close contacts. When information on a job opportunity came through a network contact, it was usually a distant contact [19:26]. The literature does not address if this concept applies across ethnic cultures.

Formally, Granovetter develops weak ties by considering two actors in a network, i and $q, \in S = \{j, l, m, \dots\}$ of all actors with ties to either or both i and q. Granovetter offers the following theory relating dyadic ties to larger structures:

Theorem 1: Dyadic Tie Theory "The stronger the tie between i and q, the larger the proportion of individuals in S to whom they will both be tied, that is, connected by a weak or strong tie" [33:1360].

In other words, the amount of overlap that a given i and q have within their respective social circles dictates the strength of tie from i or q to some j in those circles. Using Figure 2.2a as an example, it is seen that when overlap is minimal, the tie is absent between i and j. When overlap is high, the tie between i and j is strong, and finally when overlap is intermediate, then the tie between i and j is considered weak.

Thus Burt utilizes this theory in that the available pool of resources are more constrained because of a limited willingness of other actors to help; however, weak ties are more likely to connect people from different social circles. Strong ties are based on close trust relationships such as familial and close friends. Van der Hulst offers the graphic in Figure 2.3 indicating how the strength of ties are related to resources.

Figure 2.3 emphasizes the chances for gain that are created by the ability to bridge holes between agents that are not connected [19] and shows how the role of the individual acting as a bridge, connecting two (or more) different networks has the

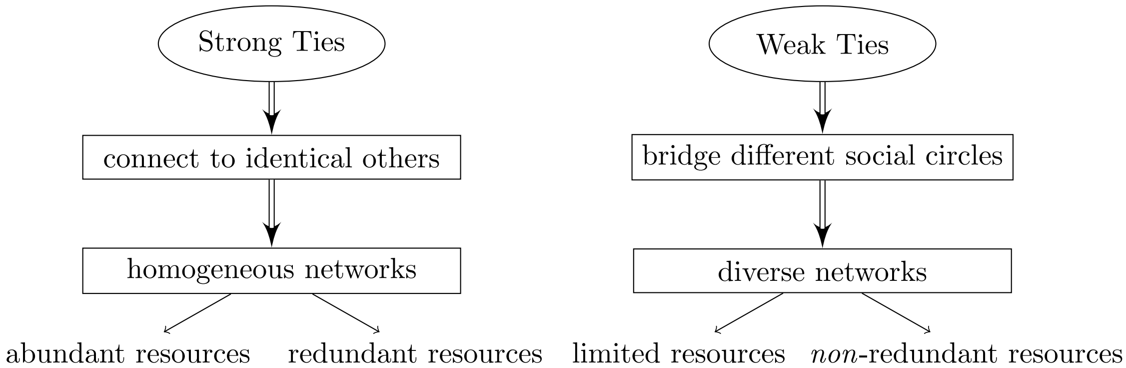


Figure 2.3 Benefits and Constraints Associated with Tie Strength

potential for high payoffs in terms of social capital. In nation building it is essential to aid in bridging structural holes across disputing political groups.

2.3.1 Structural Holes Measures.

2.3.1.1 Effective Size. Burt defines the size of a network as the number of primary contacts in a network, or the number of edges within a connected graph. Burt goes on to describe the measure *effective size*, the number of non-redundant contacts [19:47] in relationship with the ego. This is calculated by Eq. (2.7) where p_{iq} is the proportion of i's network time and energy invested in the relationship with q and m_{jq} is the marginal strength of contact j's relation with contact q, as defined in Eq. (2.6b)(2.6c). If contact j is disconnected from all contacts, then the bracketed term in Eq. (2.7) equals one, indicating that j provides no non-redundant contact in the network [19:53]. Effective size can be any positive real number ($\in \mathbb{R}^+$).

$$\sum_j \left[1 - \sum_q p_{iq} m_{jq} \right], \qquad q \neq i, j \tag{2.7}$$

Note however, that Borgatti has found a lack of correlation between Burt's effective size measure employed in structural holes when compared to ego network measures. Borgatti does not explicitly state which ego network measures are used in his comparison [17]. However, Burt directly addresses this issue of reaching 2^{nd} order contacts, or friends of friends, by using indirect access to structural holes via the measure indirect constraint [20:30-31] (presented in § 2.3.1.4).

2.3.1.2 Efficiency. Efficiency is a subset of effective size in that it is a ratio of the effective size divided by the observed number of contacts, N, or degree, of ego. Formally, it is stated in Eq. (2.8). This ratio is a number from 0 to 1 where a score of 1 indicates that every contact in the network is non-redundant. An efficiency score of 0 indicates high contact redundancy and therefore low structural hole efficiency [19:53].

$$\frac{\sum_j \left[1 - \sum_q p_{iq} m_{jq}\right]}{N}, \qquad q \neq i, j \tag{2.8}$$

The numerator in Eq. (2.8) is the effective size of i's structural holes. The component p_{iq} represents the proportion of i's network that is invested in a given alter q. m_{jq} is the marginal strength of j's contact with q (where the strength of the link between j and q divided by the strongest value that j has to any node in the network). The product $p_{iq}m_{jq}$ equals 1 when i invests all its resources in q and q is also the most important contact for j. A high score indicates that i is effective in spanning structural holes because it ties to non-redundant alters (j's) [15:529]. Efficiency is able to identify individuals with a high degree, however it also highlights individuals who only have a degree of 1. Individuals will have an efficiency of 1 when there is only one contact to invest in.

Efficiency can be relatively larger than when compared to the ego network measures Borgatti refers to of centrality and betweenness centrality [17]. By design it is large to the extent that ego's alters are connecting to different third parties, explaining the likelihood of the existence of a structural hole. Efficiency highlights the importance of 2^{nd} order relations. 2^{nd} order relations refer to how an ego is connected to an alters contact, also referred to as friends of friends.

An additional measure to highlight structural holes within a network is constraint.

2.3.1.3 Network Constraint. Network constraint is a measure that highlights the lost value when holes are missing in a network. Burt uses Figure 2.4 to illustrate i's opportunities are structurally constrained to j in that q, in whom i invests a large proportion of time and energy in both j and q, has also invested heavily in j. Thus, contact j constrains opportunities in that a large amount of i's time, energy, and resources has been invested in j when j is surrounded by few structural holes of which i could benefit. There are relationships already existing with i and other alters within this figure [19:54]. Burt defines constraint as,

Definition 26 (Network Constraint): *"Network constraint is the measure of how much a manager's time and energy are connected to a single group of interconnected colleagues - which means no access to structural holes" [20:294].*

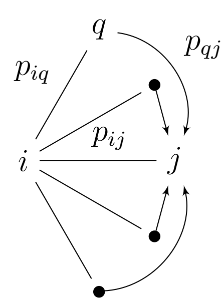


Figure 2.4 Hole Conditions of Constraint [19:52]

Constraint is defined in Eq. (2.9) where c_i is network constraint on actor i, the *ego*.

$$c_i = \sum_j c_{ij}, \quad i \neq j \tag{2.9}$$

Eq. (2.10) is a measure of c_{ij}, or actor i's dependance upon actor j [20:294]. The term p_{qj} is the proportional strength of q's relationship with j, and p_{ij} is i's relationship to j illustrated in Figure 2.4. When the product $p_{iq}p_{qj}$ is high, i's investment in q will lead back to j, indicating redundancy when p_{ij} is high [19:54]. This result also highlights a lack of a structural hole.

$$c_{ij} = \left(p_{ij} + \sum_q p_{iq}p_{qj} \right)^2, \quad q \neq i, j \tag{2.10}$$

Network constraint, as the sum of c_{ij}, measures the extent to which the manager's network of colleagues is like a straight jacket around ego, limiting their access to alternative ideas, information, and resources [20:294]. Burt describes actor j's constraint as the product of two terms: (a) the time and energy i has invested to reach j, multiplied by (b) the lack of structural holes around j [19:62].

To illustrate the calculation of this measure a simple network is presented (see Figure 2.5).

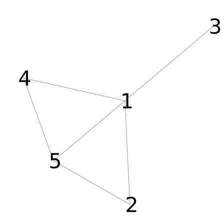


Figure 2.5 Illustrative Example of Constraint

In words, c_{ij} = Direct investment (p_{ij}) + Indirect investment ($\sum_q p_{iq}p_{qj}$). Working with actor 1 as the ego, we have the redundancy levels, or direct investment, as defined in Eq. (2.6b). The calculations for direct investment are simplified to $\mathbf{P}_{ij} = \frac{1}{n_i}$, where n_i is the degree for actor i, because the network is non-weighted and non-directed. Indirect constraint is calculated using the 2-path step distance, or $\mathbf{P} * \mathbf{P}$. An example is shown below:

$$\mathbf{P} = \begin{bmatrix} 0.00 & 0.25 & 0.25 & 0.25 & 0.25 \\ 0.50 & 0.00 & 0.00 & 0.00 & 0.50 \\ 1.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.50 & 0.00 & 0.00 & 0.00 & 0.50 \\ 0.33 & 0.33 & 0.00 & 0.33 & 0.00 \end{bmatrix} \quad \mathbf{P} * \mathbf{P} = \begin{bmatrix} & 0.083 & 0.000 & 0.083 & 0.250 \\ 0.165 & & 0.125 & 0.290 & 0.125 \\ 0.000 & 0.250 & & 0.250 & 0.250 \\ 0.165 & 0.290 & 0.125 & & 0.125 \\ 0.330 & 0.083 & 0.083 & 0.083 & \end{bmatrix}$$

Constraint between any two people then is $\mathbf{C} = (\mathbf{P} + \mathbf{P}^2)^2$, shown below:

$$\mathbf{P} + \mathbf{P}^2 = \begin{bmatrix} 0 & 0.33 & 0.25 & 0.33 & 0.5 \\ 0.67 & 0 & 0.13 & 0.29 & 0.63 \\ 1.00 & 0.25 & 0 & 0.25 & 0.25 \\ 0.67 & 0.29 & 0.13 & 0 & 0.63 \\ 0.66 & 0.41 & 0.08 & 0.41 & 0 \end{bmatrix} \text{ and } \mathbf{C} = \begin{bmatrix} 0 & 0.11 & 0.06 & 0.11 & 0.25 \\ 0.44 & 0 & 0.02 & 0.08 & 0.39 \\ 1.00 & 0.06 & 0 & 0.06 & 0.06 \\ 0.44 & 0.08 & 0.02 & 0 & 0.39 \\ 0.44 & 0.17 & 0.01 & 0.17 & 0 \end{bmatrix}$$

This example illustrates how actor 1 has the lowest constraint because of the non-redundant contact to actor 3. Constraint is found by summing across the columns in the **C** matrix above for each row, representing individual actors. Actor 3 has the highest constraint because if there were to be a relationship between actors $2, 4$, or 5, they would all be redundant with actor 1.

Borgatti offers the summary and correlation of how Burt's measures relate to social capital in Table 2.4. This table shows a positive correlation between effective size and social capital, and a negative correlation between constraint and social capital. Constraint and social capital have a negative correlation because the more an individual is constrained, the less opportunity he or she has for building social capital.

Table 2.4 Structural Hole Measures relating to Social Capital [18:31]

Name	Description	Relation to Social Capital
Effective Size	The number of alters, weighted by strength of tie, that an ego is directly connected to, minus a redundancy factor.	Positive. The more different regions for the network an actor has ties with, the greater the potential information and control benefits.
Constraint	The extent to which all of ego's relational investments directly or indirectly involve a single alter.	Negative. The more constrained the actor, the fewer the opportunities for action.

Network constraint is a measure to identify structural holes. Structural holes in a network define an actors incentive to take strategic action given certain relations, specifically to change the network in their favor. Burt offers several strategies for actors to take in order to effectively utilize structural holes, seen in Table 2.5.

Table 2.5 Burt's two sides to strategic actor strategy [19:230]

Other side: Manage constraint of absent hole	One Side: Develop the information and control benefits of an existing structural hole	
	Redundant contacts	Non-redundant contacts
Withdrawal	Withdraw from a contact in favor of his competitor	Withdraw from a contact's cluster to focus network resources in other clusters
Expansion	Add a contact's competitor to the network	Add a new cluster to the network
Embedding	Establish second relationship with contact, giving the actor more control	Establish second relationship with either or both contacts, giving the actor more control

Burt's context for these strategies is predicting the promotion rate of individuals based upon their ability to span structural holes. Of the various strategies that Burt offers in the competitive environment for business relationships, some may or may not apply to the governmental situation in a failed or failing state. Structural holes theory can initially be used to identify areas to invest time and resources in gaining social capital within the governmental structure. However, using a withdrawal strategy may not be advantageous in a government setting attempting to unite a fractured nation whereas the expansion or embedding strategies may build stronger ties.

Burt further describes the importance of building across structural holes within an organization. He states that when building a hierarchical network around a contact, the benefits are dependant upon who that network is build around [19:153], also referred to as a strategic partner. When building these networks, the greatest benefit for an individual's promotion is from building not around the immediate boss, but around a person completely removed from the immediate work group both formally and informally. By using a person completely removed from the work group, structural hole effects are most evident for those individuals on a social frontier, or where two social worlds collide [19:163]. These frontiers are most pronounced as the political boundary between top leadership and the rest of the organization [19:164]. This thesis applies this theory to investigate failed and failing states in order to gain more cooperation within the nation, ultimately leading to a viable peace.

2.3.1.4 Network Indirect Constraint. If direct contacts in an ego, i, network are the people with whom i has direct personal contact with, then the indirect contacts are contacts of contacts reached only through the direct relationships that i has with any given contact j. Burt proposes a measure to capture indirect access to structural holes through i's contacts by capturing constraint on i's contact is indirectly a constraint on i [19:30]. Thus an ego, i, with low indirect constraint has relationships with individuals who are rich in structural holes themselves, meaning that i has indirect access to these opportunities in filling the holes. This indirect access to structural holes is computed by aggregating constraint in networks around each of i's contacts, as seen in Eq. (2.11),

$$IC_i = \sum_j \delta_{ij} C_j, \quad i \neq j \tag{2.11}$$

where C_j is direct network constraint on contact j as computed in Eq. (2.9), and δ_{ij} is a weight for pooling contact networks. Burt tested several ways to compute this δ_{ij} and settled on the arithmetic average across i's contacts (i.e., $\delta_{ij} = 1/N, N =$ number of i's contacts), for a detailed discussion see [20:300-305].

Burt argues that this measure does not capture total indirect constraint, which contains the two components of 1) connections within the network around i, and 2) connections across the network around each alter j. Averaging scores across alters, IC_i captures the first component of total indirect constraint, and some unknown portion of the second component [20:301]. Burt notes the limitations to this measure when trying to evaluate 3^{rd} order contacts and higher, and can in fact be unproductive in attempting to do so. Burt illustrates this with empirical evidence given in Table 2.6. The first column is length of the path distance from i to alters j included in network around i, the second column is the indirect constraint standard deviation, and the third column is the correlation between direct and indirect constraint measures. As the network around i expands to include distant alters, or increasing the number of contacts i must go through in order to reach that alter, the correlation between direct and indirect goes to negative one, and will decrease even faster for smaller more dense networks [20:302].

These results indicate that using the indirect constraint measure provides a good description of 2^{nd} order relations (i.e., friends of friends), but not necessarily beyond that.

Table 2.6 Average Constraint from Increasingly Distant Alters [20:302]

Maximum Path Distance to Averaged Alters	Standard Deviation in Indirect Constraint	Correlation between Direct and Indirect Constraint
1	4.68	0.31
2	1.51	-0.09
3	0.92	-0.52
4	0.25	-0.76
5	0.11	-1.00

2.4 Statistical Testing

Within the field of statistical testing, there are experiments that necessitate the ordering or ranking of measures. These experiments are of special interest in the field of social science. The use of nonparametric statistical methods are useful for analyzing this type of data [63:741]. In particular, rank tests can be applied to all types of populations: continuous, discrete, or mixtures of the two [24:214-215]. These methods are also useful in making inferences in situations where doubt exists about underlying assumptions within standard statistical analysis. Wackerly, Mendenhall, and Scheaffer say there is no standard definition of nonparametric statistics, but they offer the following,

Definition 27 (Parametric statistics): *"parametric methods are those that apply to problems where the distribution from which the sample is taken is specified except for the values of a finite number of parameters. Nonparametric methods apply in all other instances" [63:742].*

Wackerly *et al.* states that nonparametric methods often are more powerful in detecting population differences when the assumptions are not satisfied when compared to standard parametric methods [63:742]. Conover describes several instances when ranks may be considered preferable to actual data. First, if the numbers assigned to observations have no meaning by themselves. Second, even if the numbers have meaning but the distribution function for the population is not a normal distribution function [24:215].

This thesis employs the use of nonparametric statistics in order to detect differences between measures of networks. Another limiting factor to using parametric statistics is the assumption of independence cannot be applied SNA, as all data is dependent upon one another [60:109].

2.4.1 Rank Correlation Coefficient. According to Conover, a measure of correlation is a random variable that is used in situations where the data consists of pairs of numbers, i.e. $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ [24:250], using (X, Y) when referring to (x_i, y_i) in general. Conover offers the following conditions for the measure of correlation to be acceptable [24:250]:

1. The measure of correlation should assume only values between -1 and $+1$.
2. If larger values of X tend to be paired with larger values of Y and smaller values of X are paired with smaller values of Y, then the measure of correlation should be positive (close to $+1$ if tendency is strong).
3. If larger values of X tend to be paired with smaller values of Y and smaller values of X are paired with larger values of Y, then the measure of correlation should be negative (close to -1 if tendency is strong).
4. If values are randomly paired with one another, then the measure of correlation should be close to zero.

Spearman's ρ_s is often used as a test statistic to test for independence between two random variables, and is insensitive to some types of dependence [24:254]. To detect correlation between two ranked variables Spearman's correlation coefficient, ρ_s is used. The test statistic tests the hypothesis of no association between two populations. There are a two assumptions for Spearman's ρ_s:

- The n pairs of observations (x_i, y_i) have been randomly selected from their respective populations, implying random assignment of n ranks within each sample [63:785].
- Each random assignment for two samples represents a sample point associated with the experiment and ρ_s can be calculated for each [63:785].

For this test statistic, the variables of interest are the ranks of each measure. Ranks for tied observations are obtained by averaging the ranks that the tied observations would occupy [63:784]. The test statistic is seen in Eq. (2.12a). When there are no ties in either the x or y observations, the equation simplifies to Eq. (2.12b).

$$\rho_s = \frac{\sum_{i=1}^{n} R(x_i)R(y_i) - \frac{1}{n}\left[\sum_{i=1}^{n} R(x_i)\right]\left[\sum_{i=1}^{n} R(y_i)\right]}{\sqrt{\left\{\sum_{i=1}^{n}[R(x_i)]^2 - \frac{1}{n}\left[\sum_{i=1}^{n} R(x_i)\right]^2\right\}\left\{\sum_{i=1}^{n}[R(y_i)]^2 - \frac{1}{n}\left[\sum_{i=1}^{n} R(y_i)\right]^2\right\}}} \tag{2.12a}$$

$$\rho_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)} \tag{2.12b}$$

Where $d_i = R(x_i) - R(y_i)$ and $R(x_i), R(y_i)$ are the corresponding ranks for paired observations (x_i, y_i).

The hypothesis for this test, H_0, is that there is no association between the rank pairs, or the pairs are mutually independent. The alternative hypothesis, H_1, asserts there is an association between the rank pairs; either a positive correlation or a negative correlation necessitating a two-tailed test. The rejection region for a two tailed test includes values of ρ_s near $+1$ and near -1. This rejection is based upon critical values, ρ_0 found in Table B.1. Reject H_0 if $\rho_s \geq \rho_0$ [63:786].

It should be noted there are other measures for tests of correlation, namely Kendall's τ. Conover states that Spearman's ρ_s tends to be a bit larger in absolute value than Kendall's τ. However, there is no reason to prefer one test over the other based upon significance because both will produce nearly identical results [24:258]. Spearman's ρ_s is employed in this thesis because Kendall's τ is usually considered to be more difficult to compute than Spearman's ρ_s [24:256].

The next section explores how these structural hole measures are applied in business relationships.

2.5 SNA Literature for Building Business via Structural Holes

Structural holes theory has been applied to business models and strategic partnerships since the early 1990's. One example is Berardo's article on how informal collaboration effects organizational success [15]. Fundamental to Organizational Network Analysis (ONA) is the idea that people work well when they work together, and even more so when the right people are connected together. Berardo tested the expectation that an organization, specifically a business organization seeking profit in dealing with a local state government, can more easily reach its goals when it taps other actors in the network for *useful* resources [15:522]. His findings reveal that, in the business arena,

> "Increasing the number of partners in collaborative practices is beneficial as long as it is guaranteed that the partners will not add unmanageable complexity by feeding new resources that the lead agency is not prepared to process effectively" [15:535].

This result validates Burt's structural hole theory to the extent that the new capacity brought to the table by bridging these holes does not overwhelm the business with infinite possibilities. Thus, a balance exists between maximizing informal collaboration while minimizing the new learning curve of information and possibilities.

Western, Stimson, Baum and Gellecum extended social capital theory by developing a set of indicators focused on the measurement of social capital [67]. They extended the degree to which social capital is related to quality of life based on case studies in Australia [67]. Western *et al.* found that there was a high correlation of quality of life with social capital [67:1106-1107], indicating the importance of informal networks in building communities.

Finally, Walker, Kogut, and Shan demonstrate that in biotechnology, network formation and industry growth are influenced by the development and maintenance of social capital [64:109]. Walker *et al.* discuss that building interfirm ties create the structure that allows for new cooperation between those firms [64:118] emphasizing the importance of building relationships with those actors who fill structural holes, thus building social capital [64:122].

2.6 Failed and Failing States

The United States and her allies have a long history of aiding the building and reconstruction of failed and failing states. JP 3-24 COIN provides the following definitions for failed and failing states:

Definition 28 (Failed State): *"A failed state may only have remnants of a government due to collapse or regime change or it may have a government that exerts weak governance in all or large portions of its territory. A failed state is unable to effectively protect and govern the population. A failed state may not have a national government with which to work and, consequently, conducting COIN is difficult, especially with respect to legitimacy at the national level. Under these extreme circumstances, the intervening authority has a legal and moral responsibility to install a transitional military authority" [42:I-3].*

Definition 29 (Failing State): *"The failing state is still viable, but it has a reduced capability and capacity to protect and govern the population. When a state is fighting an insurgency and its ability to protect and govern the population starts to decline, the pace of that states decline tends to accelerate towards collapse. Outside support for a failing states COIN efforts may halt and reverse this trend; however, assistance becomes more difficult based on the level of decline at the time of intervention" [42:I-3].*

JP 3-24 continues,

Definition 30 (Recovering State): *"The recovering state is moving towards normalcy but may have an imperfect level of viability. This state is able to protect and govern its population to some degree. A key consideration is whether the population considers the level of protection and governance acceptable and normal. A recovering state may still suffer from insurgency, although any insurgency in a recovering state will be relatively weak. When dealing with a recovering state, US efforts focus on building host nation capability and capacity and preventing a latent insurgency from emerging" [42:I-4].*

Covey *et al.* have extensive experience with U.S. operations in Kosovo, an ethnically severed and failing state, where lessons learned can be applied to the pursuit of a viable peace within failed and failing states [25]. Each case of statehood is unique; however the quest for viable peace across states is not unique. Covey *et al.* emphasize the issue of conflict while reconstructing. If conflict within the failed state has been removed, then the strategy can be focused purely upon reconstruction, however if conflict has not been removed, then the strategic focus must be "conflict transformation" [25:6]. Covey *et al.* go on to assert,

> "Rebuilding efforts alone cannot extinguish conflicts that continue to smolder or transform extremist power structures that copiously fuel the fires . . . building peace at the end of war is not a straightforward matter of handing power over to local leaders . . . , A strategy to transform internal conflict, coupled with long-term reconstruction efforts is the only realistic approach for policy makers and practitioners" [25:7].

This conflict transformation must take place for a viable peace to emerge from intolerant, no win confrontations to a system of governance where the balance of power is conducted through nonviolent means [25:9].

However, in terms of the international community aiding with the failed or failing state, if the host nation's capacity to rule the country is insufficient then, the international community will have to intervene on an interim basis rooted in relations between international and local actors to retrain and reestablish the legal system [25:11]. This concept is strikingly similar to locating and filling structural holes in order to build capacity within the host nation in order to reach a viable peace.

Covey *et al.* emphasize the importance of the custodian of the peace process, or the central actor responsible for achieving viable peace within a failed or failing state. Success hinges upon this custodian and how, "adroitly this appointed leader mobilizes international support and unifies the various components of the mission behind realistic strategies to implement the peace process" [25:17-18]. It is the premise of this thesis that these implementations can be aided via SNA and judiciously filling structural holes to mobilize aspects of the peace process in failed or failing states.

Ghani and Lockhart propose that there is a gap, the "sovereignty gap", between the custodian of the peace process and the international system in place to aid these failing states [31:3]. They go on to say that the "failed state is at the heart of a worldwide systemic crisis that constitutes the most serious challenge to global stability" [31:4]. The framework they propose defines the functions of the state, the structure to perform those functions, and networking actors. They argue for a citizen-based approach with a compact between citizen, state and the market, versus that of a top-down hierarchical imposition of the state [31:7]. This structure is strikingly similar to informal networks and the concept of social capital applied to a state level. Ghani and Lockhart emphasize the importance of positive relations between citizens and the state, and between the state with the international community for the expressed purpose to empower and involve citizens in decision making regarding resources to ensure that citizens play an active role and add value to the nation [31:8].

Ghani and Lockhart's framework contains ten functions that the state must fulfill [31:124-166]:

1. The Rule of Law. The "glue" that binds all aspects of the state, economy, and society that accounts for decision rights, processes, accountability, freedoms, and duties within the state [31:125].

2. A Monopoly on the Legitimate Means of Violence. A state must control the use of violence within it's boarders to limit violence, establish legitimacy to subordinate violence, and to use calibrated force against those who threaten the state's legitimacy [31:128]
3. Administrative Control. Achieving control through hierarchical divisions in order to deliver value to the public [31:131-132]
4. Sound Management of Public Finances. Sound management is defined as the "efficient collection and allocation of resources among contending priorities that turn ideas and aspirations into concrete outcomes" [31:135].
5. Investments in Human Capital. This includes priority investment in education and health.
6. Creation of Citizenship Rights Through Social Policy. These include rights that transcends gender, ethnicity, race, class, spatial location, and religion. This build national unity and shared beliefs [31:144].
7. Provision of Infrastructure Services. A state's ability to provide security, administration, investment in human capital, and strong economy are tied with adequate transportation, power, water, communications, and pipelines [31:147].
8. Formation of a Market. State support is crucial in order to set and enforce rules, support private companies, and intervene when market fails [31:149-150].
9. Management of Public Assets. It is the state's role to ensure upkeep and use of land, equipment, buildings, cultural heritage, forests, rivers, seas, and so fourth [31:156].
10. Effective Public Borrowing. Argues for the establishment of the central bank so that public bonds can transform public savings into capital for the state [31:160].

They argue that historically, states that have performed these ten functions create synergy and opportunity for the citizen. This framework builds on the concept that the state is there to serve and provide opportunity for the citizen [31:163].

However, a state does not need to be centralized in order to be effective. Ghani and Lockhart state, "effectiveness is derived from a delineation of governance processes that assigns decision rights to the appropriate level of government" [31:165].

Finally, they remark that,

> "Security will not be guaranteed by the use of force, though military intervention might be called upon from time to time. Security will come through the creation of functioning states, whereby the failure of politics and aid is overcome by a double compact that binds citizens, their governments, and international players in webs of rights and obligations" [31:221].

emphasizing the importance of understanding the *network structure* of a society in order to ensure that these "compacts" can be made between the appropriate individuals and groups for the effective building of a nation.

In addition, USAID's fragile states strategy highlights the importance of governance, particularly that of weak governance within a failed or failing state,

> "Weak governance, particularly in the context of a country in transition is usually at the heart of fragility. However agency resources going to fragile states mostly address symptoms of fragility such as famine and humanitarian crises, instead of the source, such as weak governance" [51:17].

Changing perspective from DoS to DoD, published doctrine of how the military builds failed and failing states in the midst of severe internal conflict of the host nation will next be briefly reviewed.

2.7 COIN

To first clarify Counter Insurgency (COIN) as to counter an insurgency, insurgency must be defined. In essence, an insurgency is

> "an armed challenge to a government, from within its jurisdiction that seeks and capitalizes on the support of important segments of the population ... an attempt to win the people's allegiance not through lawful, peaceful means but through a combination of fear and promise" [32:2]

The Army Field manual 3-24 offers this definition

> "an organized, protracted politico-military struggle designed to weaken the control and legitimacy of an established government, occupying power, or other political authority while increasing insurgent control" [38:1-2].

An insurgency seeks to undermine the government by several different means in order to discredit the governing authority of the host nation.

The Joint Publication on COIN states that the "population is the critical dimension of successful COIN" [42:III-1], and also states that civilian agencies should lead the COIN efforts [42:III-2]. COIN then is "a government's [or other political authority] effort to keep the population from bowing to the fears or embracing the promises of the insurgents" [32:2], which often includes foreign backing from the international community. COIN requires the military to think like an insurgent, establish a presence within the populace, and to gain credibility within the population for the host nation's legitimacy and capabilities [42:III-1,3]. COIN, from a U.S. perspective, is to 1) produce an outcome that advances U.S. and coalition interests, and 2) to leave in place a state that is worthy of and acceptable to its people and thus less susceptible to insurgency [32:4]. In other words, to leave a state that is capable of achieving a viable peace that is on the road to a self-sustaining peace [25].

Interestingly, the RAND report on civilian COIN says that "when people look to entities other than central government for essential functions, unofficial authorities (e.g., tribal and village elders) may be the best bulwark [defense] against insurgency" [32:10]. Gompert *et al.* makes the point that in highly tribal regions in such countries suspectable to be failed or failing, the central government can be viewed as an outsider, even a foreigner. This highlights the importance of how the supporting nations and international community define success. JP 3-24 defines success of COIN as, "the isolation of insurgents from the population, and this isolation is maintained by, with, and through the population - not forced upon the population" [42:III-4].

Gompert *et al.* makes the claim that the U.S. has demonstrated weak points in reconstruction, development, capacity building, and reform. These weak points are highlighted by the U.S.'s focus on fighting insurgents with military force versus a Department of State focus of reconstruction and development in the host nation's government. The U.S. has in the past had a strong tendency to combat insurgents with force due to the fact that the military aspect of the United States is the most equipped and trained to execute this mission. The weak point that Gompert *et al.* is alluding to is the fact that issues of the Department of State are being executed by members in the Department of Defense [32:11], despite the fact that doctrine states that "civilian agencies should lead US efforts" [42:I-2].

Thus there is a need to be able to build the host nations capacity both in DoS issues and DoD issues. This thesis identifies structural holes within the government and unofficial authorities. These structural holes can then be exploited to combat, and ultimately defeat, insurgencies and to obtain a viable peace.

2.8 Summary

In this chapter, a review of SNA literature and definitions have been presented. Structural holes theory was introduced and how that theory has been applied to building business. Moreover, how structural holes can be applied to a national level in the midst of COIN operations. The next chapter will overview the methodology of implementing structural holes to the analysis of post conflict failed and failing states.

3. *Methodology*

3.1 *Overview*

The methodology presented in this chapter extends SNA techniques in order to identify structural holes within failed and failing states. Moreover, those actors that span the structural holes within failed and failing states are identified. Since humans operate in relational contexts, SNA tools provide useful insights in developing governing authority and establishing security within these states. Section 3.2.1 introduces the notation used to describe SNA. Section 3.3 defines the parameters used for data collection within this study. Furthermore, it describes the structure of the database used to manage the data collected. Finally section 3.4 briefly describes the system used to analyze the data in this study. Additionally, the methodology is applied to a test data set to illustrate its use.

3.2 *Notation*

Two types of notations are adopted: graph theory and sociometric. Graph theory refers to a *graph* with *nodes* that are joined by lines called *edges*. Sociometric notation refers to *sociomatricies* with rows and columns making up the dyadic pairs within the network. The sociomatricies and the *adjacency matricies* from graph theory are directly related to one another [65:71].

3.2.1 *Graph Theory Notation.*

Define a set of actors $\mathscr{N}$ containing g actors given as $\mathscr{N} = \{n_1, n_2, \ldots, n_g\}$. This set is a collection of nodes of a given graph $\mathscr{G}$. An actor relating to another (i.e. n_i relates to n_j or not at all) is represented as a tie between n_i and n_j with $n_i \rightarrow n_j$, or in sociomatrix form n_{ij} where order matters. This ordered set of lines representing relations is defined as the set $\mathscr{L}$. Thus, if a graph exists with nodes $\mathscr{N}$ and lines $\mathscr{L}$, it is represented as the set $(\mathscr{N}, \mathscr{L})$, on graph $\mathscr{G}$ [65:72], or $\mathscr{G}(\mathscr{N}, \mathscr{L})$.

Multiple relations are represented by $\mathscr{R}$ on a given set of actors from $\mathscr{N}$. This is represented as n_{ijr}, highlighting the relation from i to j for relation type r. For example, a single actor can have a sociomatrix for the schools he or she attended as a child, a separate one representing the tribe they belong to, and perhaps a third representing which faction of the mujahideen they fought.

3.3 Collected Data

3.3.1 SNA Boundary Specification. Population data collection for this study focuses entirely upon the governmental personnel of the failed nation-state, and more specifically, on individuals that can be found and profiled via open-source resources available through the internet.

Appropriate bounds on the data must be identified. First, consider the determination of the vertical boundary. Profiling identifies individuals in high-ranking positions in the government starting with the president, the legislature, and the judiciary. Each ministry's organizational structure is examined in as much detail as possible. These vertical bounds have been chosen primarily due to data availability constraints. Data sources include websites such as complexoperations.org and governmental websites giving organizational hierarchies, including the names of individuals filling those positions.

Consider the determination of the horizontal boundary. Defined as official governmental organizations as sanctioned by the constitution of the state, these organizations must be led by the nation state and not by outside international supporters. While nation building involves more than just the local government, the research presented in this thesis focuses on how the U.S. helps to build a viable peace through the construction of an effective self-governing state, as defined by the people of that state.

3.3.2 Data Collection. Subject based data mining initially focuses on an individual or other data that is considered (based on specific information) to be of interest [50:17]. The goal is to determine what other persons, financial transactions, movements, or interactions (e.g. a phone call, meeting, appear together in a news article, are seen in public and so forth) are related to the initial data [50:19,185-192]. Beginning with a name, *ego*, multiple open sources were used to collect data in order to identify alters thought to be associated with the ego of interest.

Selection criterion for the initiating ego in the demonstration presented in Chapter 3 was based on three items. First, if an individual is listed as being in a government position such as president, minister, governor, cabinet member, or a leader within a government ministry. Secondly if there are documented relationships to those government officials, particularly kinship and business relationships. Finally, the data collected is based off of the 26 January 2011 parliamentary inauguration [56]. A listing of government officials, and their approval status by the Parliament is in Table B.2.

Data was collected via open sources (sources are listed in the archival data records described in Section 2.2.5.1). Primary data was collected on individuals from government websites to include biographical information, education, and so forth. A complete list of websites used to collect data can be found at Appendix B. A listing of metadata searched for is found below:

- Associations with other nodes: A linkage of who interacts with whom.
- Interaction: A measure of how often different nodes have been seen or listed together in a given search.
- Degree of association: Differentiating the tie listed above as one of influence, advice, business transaction, or social interactions such as parties.
- Attributes of individuals
  - Tribe Name

- Tools or professional skill sets: This would include if they have been trained as a doctor, a lawyer, a business man, religious education, or any specific construction skills such as masonry, carpenter, and so forth. Primarily collected based upon educational degree.
- Familial relations: This includes not only immediate family but also extended family to uncles, aunts, cousins, grandparents, and so forth.
- Aliases and different spelling of names: Especially important in the translation differences from one language to the next. There is a need to capture information on the different ways a name is conveyed in different languages to ensure discussion is captured about the same individual.
- Education: Who taught or instructed this individual? What school did they attend?
- Location: Within the last year, where the individual resides for the majority of the time.

Data was collected by hand. Primarily, data for this study was from the internet, news articles, published documents, and governmental websites illustrating relationships and links within the government. Unfortunately, there is no way of knowing the complete picture of relationships as the archival data is widely disbursed and often conflicts with other posted data. In order to mitigate conflicting information, official sources such as the most current government sponsored websites and official documents were used to the greatest extent possible. Because of these limitations, the analysis presented in Chapter 4 should be considered illustrative of how the approach could be applied with a more complete dataset.

3.3.3 Database Structure. All data was assembled into a Microsoft Access© database. Several tables are broken out in order to capture the metadata from the list found in § 3.3.2. Each record within the table captures the information pertaining only to the actor attributes within that given table. This data set has four primary tables of input, *Name, Group, Relationship* and *Affiliation.*

The name table contains attributes directly associated with the actor, e.g., the name table contains fields such as first, middle, and last names, birth place, what languages they speak, and so forth. The group table contains information about formal groups, e.g., the Ministry of Commerce, the Judicial Branch, the Presidential cabinet, what known tribes exist, ethnicity, and political parties. The relationship table contains different types of relationships possible, e.g., acquaintance, family, secretary, teacher, and so forth. This table can be directional to show direction of relationship, for example, a father has a directed fatherly relationship with his son. Finally, the affiliation table contains what sort of affiliation an actor can have a with a particular group. In order to capture the degree of involvement an actor can have in that particular group an edge weight was added to the association measure. However, since there is no published standard for weights of relationships, this data is not included in this thesis but would be an area for further research. Refer to Tables 3.1 and 3.2 for a complete listing of these four tables.

Table 3.1 Name and Organization Data Tables

Names
First
Middle
Surname
Title
Birth Place
Languages
Email

Group & Organization
Tribe
Ethnicity
Formal Group

Each table serves as a basis in building the edge lists for analysis. The relationships are captured in Figure 3.1. Analysis begins with the edge lists. The primary edge lists are listed in the middle of the figure where the relation matrix, capturing who is related to whom, is read into SNA analysis software as a directional graph

Table 3.2 Relationship and Affiliation Data Tables

Relationships		**Affiliation**
Type	Directed	Type
Acquaintance		President
Business		Minister
Cousin		Director
Daughter	✓	Governor
Father	✓	Leader
Granddaughter	✓	Deputy
Grandfather	✓	Member
Grandson	✓	Political Supporter
Mother	✓	Financial Supporter
Secretary of		Sympathetic
Sibling		Neutral
Son	✓	Disgruntled
Spouse		Opposition Supporter
Teacher	✓	Actively Opposed

$\mathscr{G}$. The affiliation matrix associates individuals with different organizational groups, ethnicity, tribes, and professional groups (e.g., a doctor in America could belong to the American Academy of Pediatrics). The third matrix in the center of the figure is the education matrix containing who has attended what school, from elementary through graduate school. Analysis in this thesis focuses primarily on these three relational matrices, and is read into SNA analysis software as node edge lists.

The next section will describe the software used to perform network analysis.

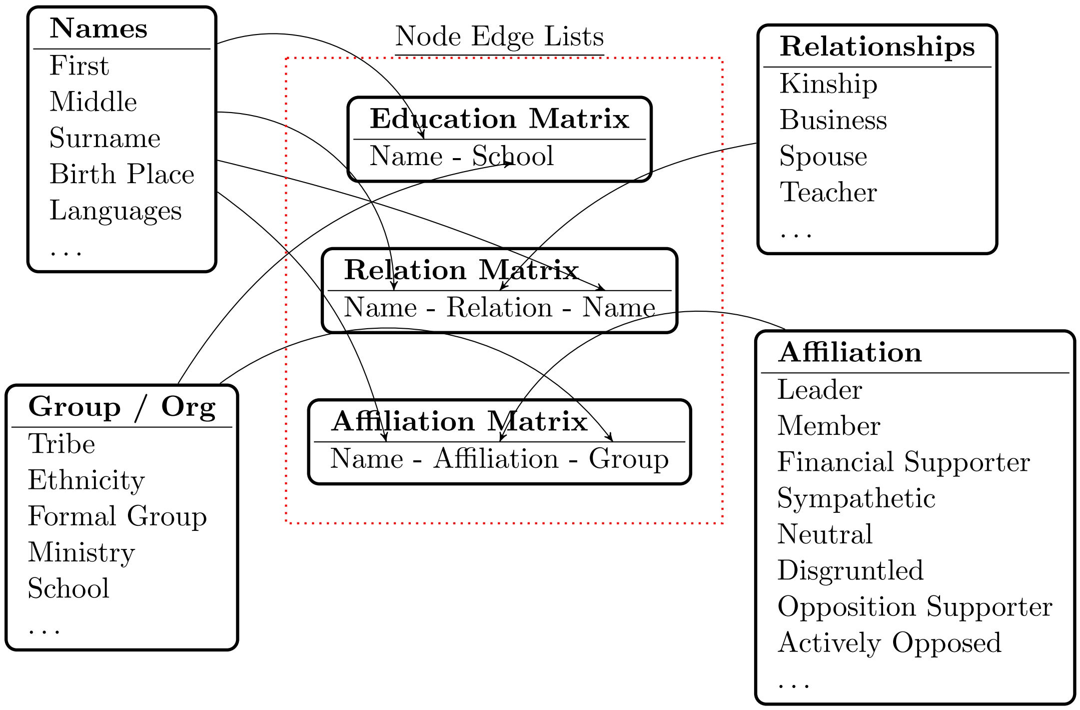


Figure 3.1 Access Database Data Structure

3.4 Coding

3.4.1 Python. Python developers state Python is a dynamic object-oriented interactive high-level programming language comparable to C^+, Visual Basic, Java, and so forth. Python is open source, and free to use, modify, and redistribute because of the OSI-approved open source license and can thus be used for commercial products without limitation [7] . Python is a notable language because it:

- Uses easy to use language ideal for prototype development and ad-hoc programming tasks [7].
- Contains large standard library supporting many common tasks like searching text, connecting to databases, reading/writing files [7].
- Cross platform capability that is easily embedded within an application providing a programmable interface [7].

- Free software, meaning no fee to download or use, and can also be freely modified and re-distributed [7].

Experience has shown that Python is a viable approach to allow for users across platforms to be highly productive across multiple disciplines (e.g., communication, social, data and biological networks) [35]. One of the many advantages of Python is its modular coding language structure allows users to create and define their own modules. This allows the creation of specific programming with portable files, which is at the core of analysis in this thesis. This thesis employs Python version 2.7.

3.4.1.1 NetworkX. NetworkX is a Python module package designed specifically for the creation, manipulation, and study of the structure, dynamics and functions of complex graphs and networks [35]. NetworkX was originally inspired by Guido van Rossum's Python graph representation essay [62] and it was developed and maintained as an open source software package at Los Alamos National Laboratory by Aric Hagberg. Hagberg states NetworkX was designed for mathematicians, physicists, biologists, computer scientists, and social scientists [35] as a tool to study the structure and dynamics of social, biological and infrastructure networks.

Hagberg maintains that NetworkX allows for fast analysis of graphs and is limited only by the size of the computer, not the software. Thus, there are virtually no limitations on size or structure of the network beyond platform specifics. This enables NetworkX to handle large, nonstandard data sets [35]. It is advantageous when using SNA software to have the flexibility to interface with differently structured databases. NetworkX employs a graph adjacency list representation based on the Python dictionary data structure. This structure allows for fast addition, deletion, and lookup of nodes and neighbors in large graphs. This thesis employs NetworkX 1.3.

NetworkX is highly documented and linked to the peer-reviewed articles used to create the algorithms within the code [36]. This code is freely available to use, modify, and re-distribute. Burt's structural holes measures, originally coded in *structure* [19:50], has been coded in several other SNA software packages including iGraph [54], Pajek [14], as well as in NetworkX. In order to validate the structural holes algorithms, a random graph was created and analyzed in the following section.

3.4.2 Code Validation. The structural holes algorithm was coded by Diedrik van Liere for NetworkX and modified slightly by Aric Hagberg [61]. This code can be found in Appendix A.1. In order to validate the code and demonstrate analysis performed in this thesis, a test data set was randomly created. It can be found in Appendix B.1 and is graphically represented in Figure 3.2. Note this example data set has some interesting properties in that there are some highly connected actors, distinctly different groups, individuals spanning between different groups, and individuals with minimal connections.

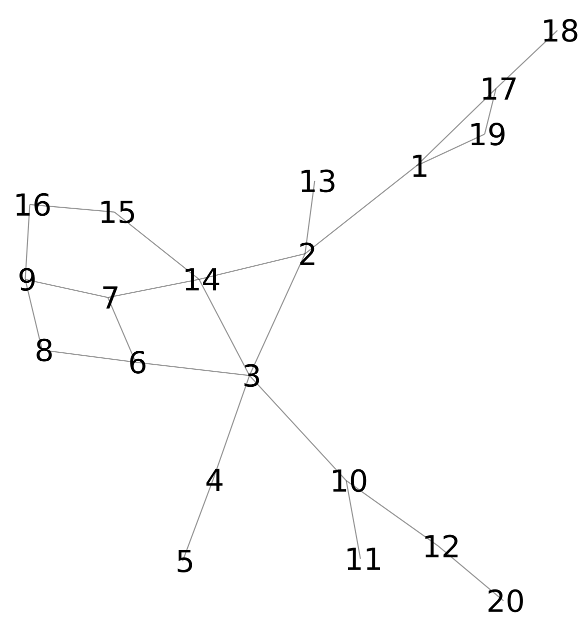


Figure 3.2 Graph of Test Data Set

3.4.2.1 Ego Measure Analysis. In this section, the data set is analyzed by applying the structural hole measures of effective size (see Eq. (2.7)), efficiency (see Eq. (2.8)), network constraint (see Eq. (2.9)), and indirect network constraint (see Eq. (2.11). Burt's measures are used in conjunction with other ego network measures such as centrality (see Eq. (2.1)), betweenness centrality (see Eq. (2.3)), and eigenvector centrality (see Eq. (2.5)). Finally, a definition and explanation of hole signature is presented.

The standard ego network measures for each node (centrality, betweenness centrality, and eigenvector centrality) are computed and presented on the left hand side of Table 3.3. If this data is sorted by centrality from highest to lowest, we would find that nodes 3, 2, and 14 all have a centrality score of over 0.20. The numeric value of 0.20 is not significant in itself, however in this example there is a significant gap between these top three to the next closest actor. This gap can be observed in Figure 3.3a, where the actors have been sorted by centrality and graphed along a number line in order to show separation, with the best scoring actors on the upper right side of the graph. These three nodes have a degree of 4 or higher. It can be observed from the graph in Figure 3.2 that without these three actors the graph would not be complete; it would fragment.

Analysis based on betweenness centrality now includes actors 1 and 10 (with values over 0.20) in addition to actors 2, 3 and 14, as graphically represented in Figure 3.3b. This highlights what the betweenness centrality measure identifies; it reflects the number of actors with whom an individual is connecting indirectly. For example, in the graph in Figure 3.2, actor 10 is connecting actors 11, 12, and 20 to the rest of the network by being connected to actor 3. In other words, 10 is acting as a bridge to the main group through 3. Burt asserts that betweenness is, "a count of the structural holes to which a person has monopoly access" [20:297]. However, Freeman developed betweenness to describe centrality in small, five person

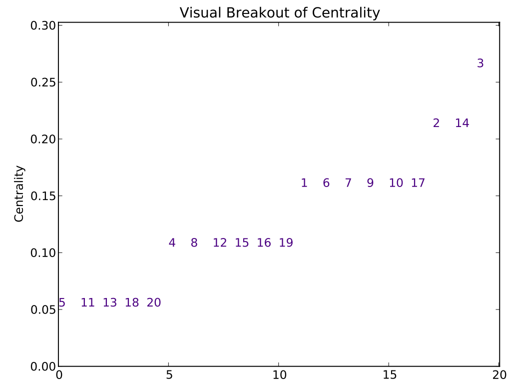
Table 3.3 Standard Ego Network Measures and Structural Hole Measures

Node	Centrality	Betweenness Centrality	Eigenvector Centrality	Effective Size	Efficiency	Direct Constraint	Indirect Constraint
1	0.16	0.28	0.19	2.33	0.78	0.56	0.59
2	0.21	0.43	0.39	3.5	0.88	0.31	0.53
3	0.26	0.58	0.48	4.6	0.92	0.25	0.36
4	0.11	0.11	0.17	2	1	0.5	0.62
5	0.05	0	0.06	1	1	1	0.5
6	0.16	0.13	0.3	3	1	0.33	0.36
7	0.16	0.08	0.29	3	1	0.33	0.33
8	0.11	0.02	0.15	2	1	0.5	0.33
9	0.16	0.03	0.18	3	1	0.33	0.44
10	0.16	0.29	0.2	3	1	0.33	0.58
11	0.05	0	0.06	1	1	1	0.33
12	0.11	0.11	0.07	2	1	0.5	0.67
13	0.05	0	0.13	1	1	1	0.31
14	0.21	0.27	0.43	3.5	0.88	0.31	0.35
15	0.11	0.08	0.18	2	1	0.5	0.41
16	0.11	0.01	0.11	2	1	0.5	0.42
17	0.16	0.11	0.1	2.33	0.78	0.56	0.82
18	0.05	0	0.03	1	1	1	0.56
19	0.11	0	0.09	1	0.5	0.89	0.56
20	0.05	0	0.02	1	1	1	0.5

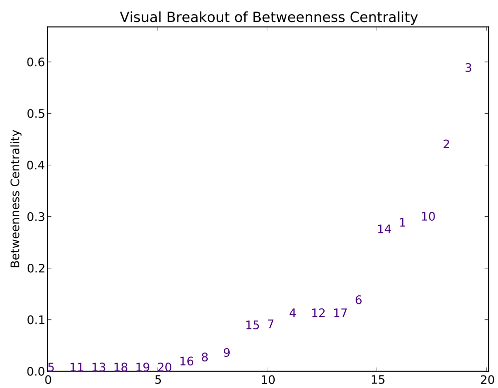
task groups in laboratory experiments and should be limited in its use for indirect contacts outside of the small group [20:299].

Eigenvector centrality has actors 3, 14, 2, 6, and 7 with the highest scoring as seen in Figure 3.3c. It is interesting that now actors 6 and 7 are considered to be central, but since eigenvector centrality takes into account who the actor is connected to, it is seen that actors 6 and 7 are connected to the key players in the network, namely 3 and 14.

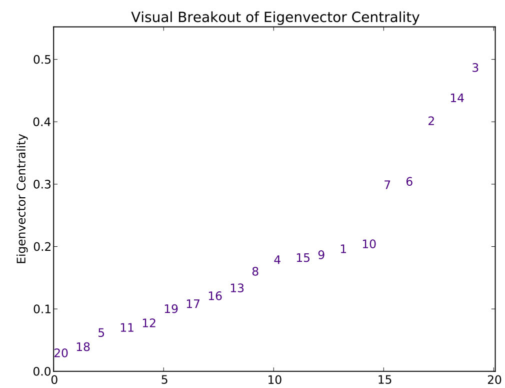
The data on the right side of Table 3.3 contains Burt's structural holes measures. Sorting the data based on effective size (see Figure 3.4a), it is seen that nodes $2, 3$, and 14 rank high, but nodes $6, 7, 9$, and 10 also have a relatively high measure of 3.0, when compared to other actors within this network. This shows that node 3 has the most non-redundant contacts within the graph, or the most contacts who



(a) Centrality



(b) Betweenness Centrality



(c) Eigenvector Centrality

Figure 3.3 Visual Breakout of Centrality, Betweenness, and Eigenvector Centrality on Test Data Set

are not in contact with one another. One point to note is that even though actors 6 and 7 are inter-related within contacts $3, 8, 9$, and 14, they still have a high effective size as they are connected to individuals who are not directly connected with one another. By counter example, node 19 only has an effective size of 1.0, even though 19 is connected to two nodes, actors 1 and 17. Since actors 1 and 17 are directly connected, the effective size of 19's network is now only one. This example highlights aspects of redundancy.

Burt asserts that dense networks are more constraining since there are more connections [20:296]. The density of this network is calculated to be 0.12, indicating that there should be a high correlation between degree and effective size.

The low density indicates that there is relatively very little redundancy in the organization due to the difference between Burt's effective size measure and the degree of the highly connected actors. This lack of difference is confirmed by the Spearman's ρ_s correlation coefficient between degree and effective size (see Table 3.4). The values for Spearman's ρ_s were calculated using the scientific tools for Python statistics package, SciPy [10]. The small p-value indicates rejection of H_0; in this network, there is no statistical difference between degree and effective size. However, if the network were a complete graph (i.e., all egos connected to all alters) then a statistically significant difference between degree and effective size would be expected due to the abundant redundancy in a complete network. Thus, effective size is a function of network density [20:298].

Analysis using Burt's efficiency measure, which is the effective size over the degree of ego (see Figure 3.4b), indicates that there are several efficient actors because there are several actors with degree of only one. This is highlighted with node 5, who has one non-redundant contact, and a degree of one, thus perfect efficiency. However, even though there is perfect efficiency, that does not mean that node 5 has access to structural holes. Examining node 19 (see Figure 3.2), it is seen that this actor has one effective contact, but a degree of two; thus efficiency is 50%. Care must be taken in analyzing efficiency within network $\mathscr{G}$.

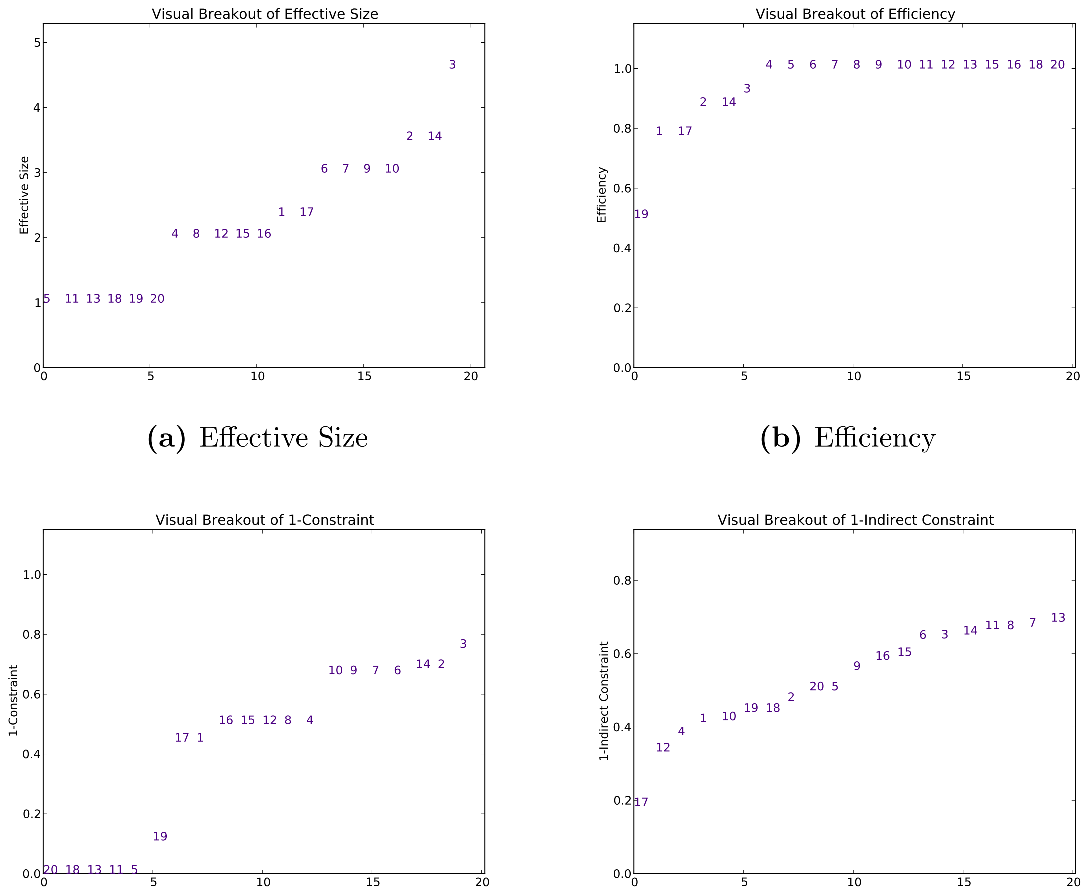
For constraint, the data needs to be sorted from lowest to highest, as constraint negatively effects social capital (see Table 2.4). The visual breakout in Figure 3.4c shows 1-c_{ij} in order to standardize the graphs so that the best scoring actor is still located in the upper right of the graph. In doing so, it is seen that actor 3 has the best score (i.e., lowest constraint). This result indicates that actor 3 is most effectively allocating resources across structural holes or gaps in the network in order to have

Table 3.4 Test Data Set Degree versus Effective Size with Spearman's ρ_s Correlation Coefficient

Node	Degree (D)	Effective Size (ES)	Difference ($D - ES$)
1	3	2.33	0.67
2	4	3.5	0.5
3	5	4.6	0.4
4	2	2	0
5	1	1	0
6	3	3	0
7	3	3	0
8	2	2	0
9	3	3	0
10	3	3	0
11	1	1	0
12	2	2	0
13	1	1	0
14	4	3.5	0.5
15	2	2	0
16	2	2	0
17	3	2.33	0.67
18	1	1	0
19	2	1	1
20	1	1	0
ρ_s	0.968		
p-value	2.6×10^{-12}		

the most access to non-redundant resources. According to Burt, "if connections with other players matter, the manner in which you are connected matters" [19:118]. Burt argues that an actor rich in structural holes, or is surrounded by structural holes, "are the actors who know about, have a hand in, and exercise control over more rewarding opportunities" [19:116] and he measures brokerage opportunities, or structural holes, with the summary index of network constraint.

Finally, indirect constraint measures an actor's access to structural holes through its contacts' access to holes, or friends of friends. The visual breakout at Figure 3.4d shows $1 - IC_{ij}$ just as constraint does. Interestingly, it is seen that actor 13 is in



(a) Effective Size **(b)** Efficiency

(c) $1-$ Direct Constraint **(d)** $1-$ Indirect Constraint

Figure 3.4 Visual Breakout of Structural Hole Measures on Test Data Set

the best position to utilize access to structural holes from actor 2's network. This is because actor 13's only contact is actor 2, who is spanning structural holes.

A summary of how each node is sorted according to each individual measure can be found in Table 3.5. In this table, each column is sorted to show the rank of the nodes as they score regarding the measure along the top. Rank is determined by score; a high score yields a better rank. For example, the three highest ranked nodes for betweenness centrality were nodes $3, 2$, and 10. Ranking for direct and indirect constraint based upon the measure being sorted from smallest, or least constrained, to highest, or most constrained. For example, the three least constrained nodes are $3, 14$, and 2, indicating that they rank the best in the constraint column.

Table 3.5 Test Data Set Nodes Sorted by Measure

Rank	Centrality	Betweenness Centrality	Eigenvector Centrality	Effective Size	Efficiency	Constraint	Indirect Constraint
1	3	3	3	3	4	3	13
2	2	2	14	14	5	14	7
3	14	10	2	2	6	2	8
4	1	1	6	10	7	10	11
5	6	14	7	9	8	9	14
6	7	6	10	7	9	7	3
7	9	17	1	6	10	6	6
8	10	4	15	17	11	16	15
9	17	12	9	1	12	15	16
10	4	7	4	16	13	12	9
11	8	15	8	15	15	8	5
12	12	9	13	12	16	4	20
13	15	8	16	8	18	17	2
14	16	16	17	4	20	1	19
15	19	19	19	20	3	19	18
16	5	5	12	19	14	20	10
17	11	11	5	18	2	18	1
18	13	13	11	13	17	13	4
19	18	18	18	11	1	11	12
20	20	20	20	5	19	5	17

Table C.1, in Appendix C, contains the Spearman correlation coefficient for all pairwise comparisons of the test measures. The table correlations are summarized in Table 3.6 representing only those measures that were significantly correlated with a p-value of < 0.05, indicating a rejection of H_0. These values are sorted by ρ_s from smallest to largest. The results indicate that effective size, eigenvector centrality, degree, centrality, and betweenness centrality are all negatively correlated with constraint, indicating that the larger the number of contacts an actor has, the less that actor is constrained by redundant relations for this example. This also leads to questions regarding the differences between the measures of effective size, eigenvector centrality, degree, centrality, and betweenness centrality. These differences, or lack of difference, in measures needs to be robustly tested across a wide array of different network types, which is beyond the scope of this thesis. The only two measures that failed to reject the null hypothesis across all measures for this network were efficiency and indirect constraint. This may be due to the fact that these measures are computing aspects different from centrality for the given actor.

Table 3.6 Significantly Correlated Spearman's ρ_s Coefficients from Table C.1

Measure	Highly Correlated with	ρ_s	p-value[a]
Constraint	Effective Size	-0.9390	0.000
Constraint	Eigenvector	-0.8920	0.000
Constraint	Degree	-0.8893	0.000
Constraint	Centrality	-0.8893	0.000
Constraint	Betweenness	-0.8151	0.000
Betweenness	Eigenvector	0.8234	0.000
Degree	Eigenvector	0.8842	0.000
Centrality	Eigenvector	0.8842	0.000
Degree	Betweenness	0.8883	0.000
Centrality	Betweenness	0.8883	0.000
Effective Size	Betweenness	0.8951	0.000
Effective Size	Eigenvector	0.9089	0.000
Effective Size	Degree	0.9685	0.000
Effective Size	Centrality	0.9685	0.000
Degree	Centrality	1.0000	0.000

[a] p-values listed were $\leq 1.58 \times 10^{-7}$, or effectively 0

In addition to sorting and rank testing the values, a key actor analysis is performed by plotting the betweenness centrality versus the eigenvector centrality. Insight is gained by examining both measures simultaneously identifying actors well positioned to connect alters to other highly connected alters. Those actors ranking high on betweenness have access to possible "monopoly opportunities for brokerage" [20:297]. A monopoly on opportunity exists if an actor knows two disconnected people, then there exists one opportunity to broker a connection between people. One would also find those actors who rank high on eigenvector centrality, a measure of who is better connected. The graph at Figure 3.5 contains the actors in the field with betweenness centrality along the horizontal axis and eigenvector centrality along the vertical axis.

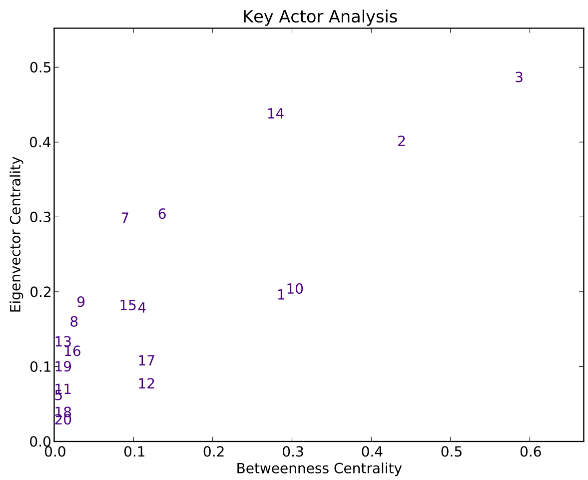


Figure 3.5 Key Actor Analysis on Test Data Set

Considering Figure 3.5, it is seen that there are several actors who are individually high scoring in either betweenness centrality or eigenvector centrality. For example, actors 1 and 10 are high with respect to betweenness and actor 14 is high regarding eigenvector centrality. Actors 2 and 3 are high in both measures as they lay in the upper right corner of the graph, along the best fit line. This result indicates that these actors (1, 10, 14, 2, and 3) are well positioned to connect unconnected individuals. For example, these are actors who are in a position to include minority groups into a government, perhaps.

Another way to visualize this same information is by a color-map graph (see Figure 3.6). This graph visually depicts the node color by betweenness and the node size by eigenvector centrality. The higher the score, the darker the color and the bigger the node respectively.

Betweenness centrality is an important measure because a structural hole is defined to occur when two people are disconnected, so betweenness is a ratio of structural holes to which an actor has direct access [20:297]. While eigenvector

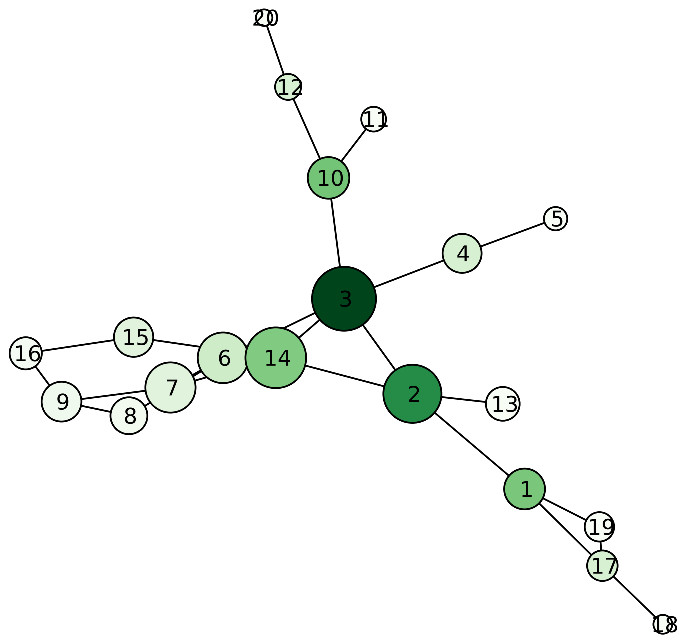


Figure 3.6 Color Map of Test Data Set
Node Color → Betweenness Centrality
Node Size → Eigenvector Centrality

centrality is a measure of who is connected to whom, it can be an indication of who has indirect access to structural holes. Thus, from Figures 3.5 and 3.6, we can see there is a substantial difference between actors 2 and 3, actors $1, 10$, and 14, compared to the rest of the actors in $\mathscr{G}$. This can be verified by performing a paired-t test assuming unequal variances between the sub-groups of actors in order to ensure there is a statistical difference between actors within each measure. Results are shown in Table 3.7. With a small $p-$value, one rejects the null hypothesis that the means of the groups are equal (i.e. $h_o : \mu_1 = \mu_2$) indicating there is a difference between the subgroups. From this data, actors $1, 2, 3, 10$, and 14 are further analyzed in the next section.

3.4.2.2 Hole Signature. Burt suggests a form of analysis called *hole signature* which compares the investment and constraint of an actors relationship

Table 3.7 Paired t-Test on Clustered Groups for Betweenness and Eigenvector Centrality

	Betweenness Centrality		Eigenvector Centrality	
	Nodes $1, 2, 3, 10, 14$	All Others	Nodes $1, 2, 3, 10, 14$	All Others
Mean	0.37	0.05	0.34	0.13
Variance	0.0181	0.0026	0.0181	0.0071
p-value	0.0062		0.0224	

to all other egos in network $\mathscr{G}$. This comparison indicates where there is much opportunity, and where there is little opportunity for the actor. According to Burt,

> "The pattern of these characteristics across relationships is a signature with which players can be identified, studied, and compared for their entrepreneurial opportunities. In the language of structural holes, the pattern is a hole signature" [19:66].

For illustrative purposes, Figure 3.7 is the hole signature for actor 1 from the network shown in Figure 3.2. The line across the top of the shaded region marked with circles describes the proportion, p_{ij}, of an actors network time and energy invested in each relationship [19:66]. The line at the bottom of the shaded region marked with triangles describes the extent to which relationship constrains the actors investment opportunities, c_{ij} [19:66]. Contacts within ego's network are listed within the field and labeled next to each point on the bottom line. The horizontal axis is a number line spanning the length of ego's contacts. In this case actor 1 has contacts with actors $2, 17$, and 19. The bottom line is close to the top line when there are few structural holes around that particular contact. Thus, actor 1 has access to many structural holes through actor 2, and has the least access to structural holes, or is most constrained by actor 17. This example is an unweighted, undirected graph, thus $p_{ij} = \frac{1}{\sum z_{ij}}$, $\forall i, j$, or is just the average degree for all contacts from i to j. If this were a weighted graph, the top line would vary accordingly to i's investment into contact j, as calculated in Eq. (2.6b) on page 2-18.

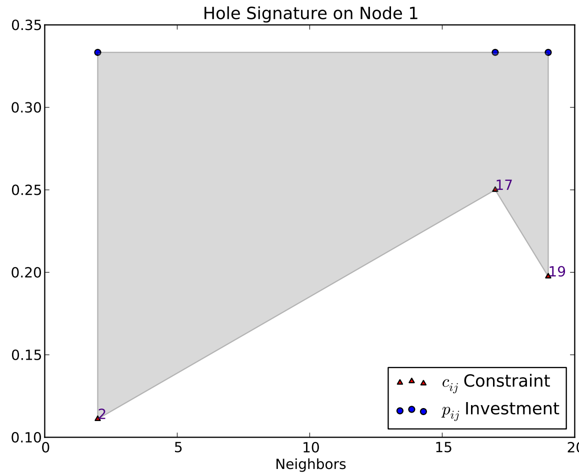


Figure 3.7 Hole Signature on Actor 1 from Figure 3.2

The shaded area is called the signature of i's ego network, which describes the opportunities and constraints across the relationships within i's network [19:66-67]. This measure is not the area under the curve, it is the vertical line from the lower point to the upper point for a given contact j, or the difference between p_{ij} and c_{ij}. The volume is standardized between zero and one in order to compare across actors. The constraint line on the bottom divides the signature into two separate regions. Below this line, the empty white space, is the constrained portion of i's interaction with j, $\sum c_{ij} = \mathbf{C}$. The shaded region above the constraint line is the unconstrained portion, $\sum p_{ij} - \sum c_{ij} = 1 - \mathbf{C}$ [19:67]. Burt asserts that the "hole signature provides a quick visual impression of the volume and locations of opportunity and constraint in a [ego's] network" [19:67]. The jagged edges of a hole signature identify relations where i has the greatest and least opportunity for spanning structural holes, and thus the greatest and least opportunity for social capital return.

This tool enables two different analytical uses, evaluation of relationship types and environmental types. Three different relationship types [19:67] are examined: opportunity, constraint, and sleeper.

Opportunity relationships are indicated by a large band of grey on the hole signature. This large band of grey is indicated by a high upper line and a low lower line in the hole signature [19:67]. In the example, actor 1 has an opportunity relationship with actor 2, and the greatest room to negotiate.

Constraint relationships are indicated by a high, narrow band in the hole signature. This type of relationship results from i investing a large amount of time in j but j is surrounded by few structural holes, and is indicated by a high upper and lower line on hole signature. In the example, Actor 1's relationship with actor 17 is considered a constrained relationship.

Sleeper relationships are ones in which i invests little time and energy in j with little to no structural holes surrounding j. Since network analysis is a snapshot in time, this is not to say that j could not become a useful asset if i's activities change in the future, making j's contacts now relevant [19:68]. The example does not contain a sleeper.

Burt identifies four environments (see Figure 3.8) that can be highlighted through hole signature analysis. The networks on the left have the corresponding hole signatures on the right. In the clique, all actors are connected with one another. In the center-periphery network, the five contacts are connected only through the actor labeled 'You'. Both have similar hole signatures, but the center-periphery network has more area under the top line, thus there are more structural holes in this network. The difference becomes negligible as the number of contacts becomes large [19:70].

The two networks on the bottom of Figure 3.8 shows hierarchy relationships. The leader is actor C, which according to Burt costs a large proportion of time and energy and yet is constrained due to the leader's connections with all other actors in the network [19:70]. Thus, networks *across* networks may span more structural holes versus networks *matching* the hierarchical structure.

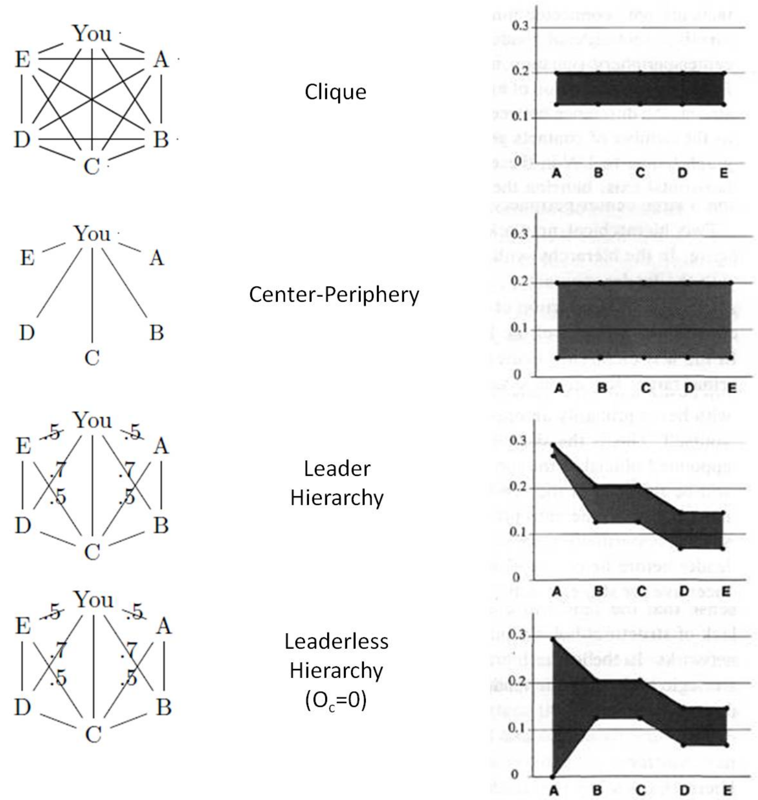
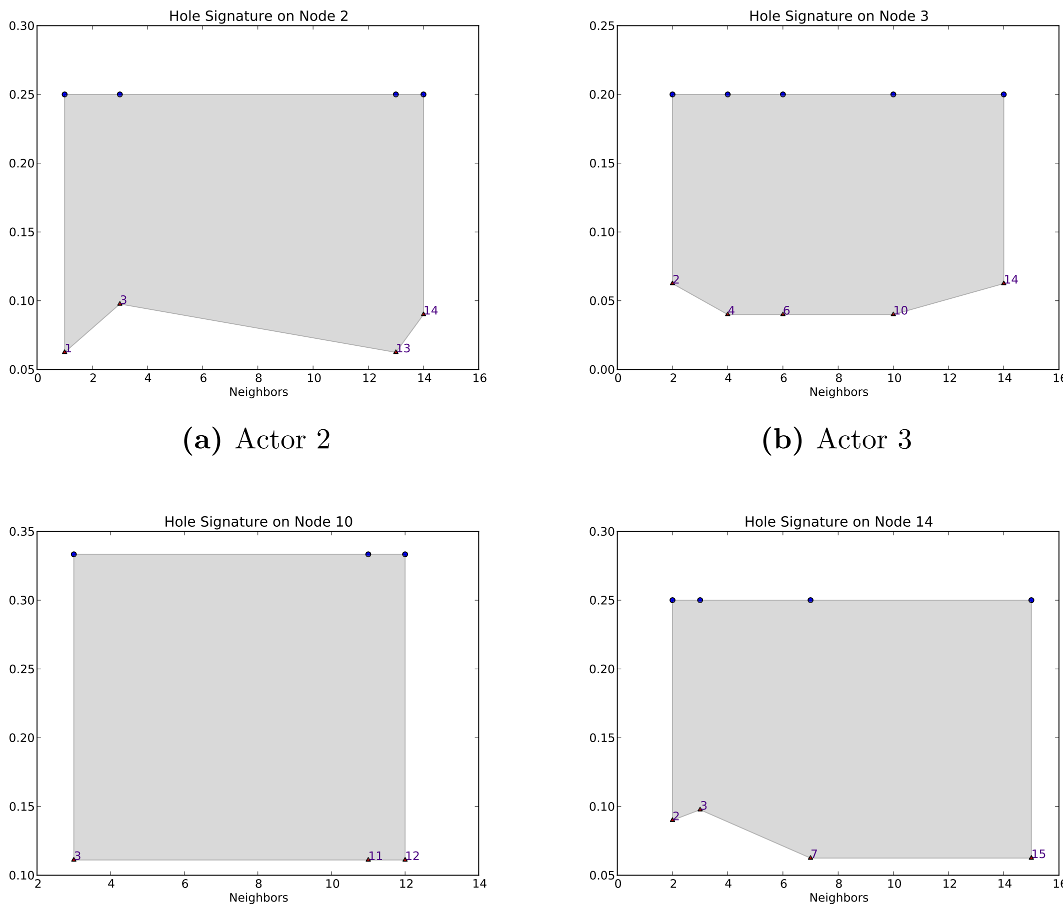


Figure 3.8 Burt's Hole Signatures for Illustrative Networks (Relations without assigned values have a value of 1) [19:69].

The use of hole signature can highlight where an individual is investing too much time with very little return in terms of spanning structural holes, access to new ideas, and social capital. Hole signature also provides insight into where an individual should spend more time investing and in what specific individual in order to span structural holes.

Referring back to the example network of Figure 3.2, hole signatures for actors $2, 3, 10$, and 14 are shown in Figure 3.9. Examining actor 2 (see Figure 3.9a), it is seen that actor 2 is constrained by redundant relationships to actors 3 and 14 by the reduced grey space at those points.

They are redundant contacts to actor 2 because actors 3 and 14 are connected with each one another, thus they constrain actor 2's access to structural holes. However, actors 1 and 13 show opportunity for spanning structural holes as they reach into subgroups that are non-redundant. In Figure 3.9c, actor 10 is an interesting actor as his hole signature is a rectangle, meaning that he has no redundant contacts thus his efficiency is 1.0.



(a) Actor 2

(b) Actor 3

(c) Actor 10

(d) Actor 14

Figure 3.9 Hole Signatures on Test Data Set for Respective Nodes

From this analysis, it is clear that actors 2 and 3 are central actors to this network and are critical players in the graph because they span several structural holes in their role as a bridge between subgroups. It can also be argued that actor 10 and actor 1 are important actors because they span into several different subgroups and act as brokers, or even middle management between leaders and workers. Burt's methodology would propose that in this example actor 2 should break ties either with actor 3 or 14 since they provide redundant contacts to the same sub-group. In doing so, actor 2 is now free to invest time and energy in other relationships that would provide for higher returns in social capital and the possibility to span structural holes. Burt shows that those spanning structural holes are in a more competitive position to accomplish their agenda, pool resources, and have a wider breadth of ideas [19, 20]. Burt's strategies are focused on identifying and leveraging causal relationships for an individual to be promoted within a given company. In the context of building a nation, these strategies of relationship withdrawal may not be optimal for SSTRO, and may be counter productive. Thus for SSTRO context, actors $2, 3$, and 14 would be individuals to empower since they are bridging these structural holes between subgroups within the network, and encouraged to continue building ties with each subgroup in hopes of strengthening relations between two subgroups.

In the next section, these analysis techniques will be applied to data collected from open source resources on failed and failing states in order to identify relationships and individuals who could potentially play a larger role in establishing the state on a path to a sustainable, viable peace.

4. Analysis, Results, and Implementation

In this chapter, the methodology presented in Chapter 3 is applied to the analysis of Afghanistan. According to the 2010 Failed States Index (FSI), Afghanistan is a failed state. This chapter is organized as follows: Section 4.1 describes the FSI, which is useful in identifying important indicators of failed states. Section 4.2 provides a brief review of Afghan history. Section 4.3 discusses limitations on the data collected for this study. Section 4.4 provides the analysis on the data set. Finally, section 4.5 grants insight regarding applications to SSTRO.

4.1 The Failed States Index

The Failed States Index is a combined effort by the Fund for Peace (FFP) [8], and the journal *Foreign Policy* [9]. The 2010 FSI can be found in Appendix Table B.4. The FSI focuses on indicators of risk based on data collected from 90,000 online English-language publications world wide, excluding social media such as blogs, twitter, facebook, and so forth [8] in order to analyze 177 countries. The FFP incorporates data from reputable institutions working in areas of nation building, comparing the data with a separate qualitative review of each country [8]. A list of the 12 indicators for a failed state, utilized by the FFP is [8]:

Social Indicators :

1. Mounting demographic pressures.
2. Massive movement of refugees or internally displaced persons creating complex humanitarian emergencies.
3. Legacy of vengeance-seeking group grievance or group paranoia.
4. Chronic and sustained human flight.

Economic Indicators :

5. Uneven economic development along group lines.

6. Sharp and/or severe economic decline.

Political Indicators :

7. Criminalization and/or delegitimization of the state.

8. Progressive deterioration of public services.

9. Suspension or arbitrary application of the rule of law and widespread violation of human rights.

10. Security apparatus operates as a "state within a state".

11. Rise of factionalized elites

12. Intervention of other states or external political actors.

Of the items listed above, this study focuses primarily on items 1, 5, and 11. An illustrative analysis of Afghanistan demonstrates the value of the research. Afghanistan is a country of high interest to the United States and its coalition partners. A brief background on Afghanistan follows.

4.2 Afghanistan - A Brief Background

The following short history is adapted from the U.S. Department of State's country study on Afghanistan [58]. Afghanistan is in a location of strategic importance bridging the middle eastern countries of Iran, Turkmenistan, Uzbekistan, and Tajikistan to the eastern Asian countries of India, Pakistan, and China. It has been refereed to as the, "crossroads of Central Asia" [58:XIII]. Ethnic turmoil has been evident in the land for centuries. Beginning in 328 BC, Alexander the Great entered the territory to capture Balkh. This precipitated invasions from peoples of a number of different backgrounds including Scythians, White Huns, Turks, and Arabs [58:XIV]. Mahmud of Ghazni turned the region into a cultural center and an operational base for raids into India around AD 1000 [58:XIV]. Genghis Khan led a Mongol invasion in 1219 slaughtering the peoples of the area, destroying fertile agricultural areas and the major cultural centers of Herat, Ghazni, and Balkh. Struggles for supremacy followed for several hundreds of years until finally, in 1747, Ahmad Shah Durrani unified the country [58:XIV]. From 1747 until 1978 all of Afghanistan's rulers were from Durrani's Pashtun tribal confederation. Rivalries between the British and Russian Empires resulted in three Anglo-Afghan wars, 1839-42, 1878-80, and 1919. All three wars are remembered for the ferocity of Afghan resistance to foreign rule [58:XV]. The third Anglo-Afghan war led to the British relinquishing their claim over Afghanistan with the Treaty of Rawalpindi in August 1919.

The 20^{th} century was a continuation of the turmoil of the previous 2000 years. King Amanullah (1919-29), while forward thinking, alienated many tribal leaders by removing the Muslim veil for women, opening co-educational schools, and undertaking substantial modernization throughout the region [58:XV]. Mohammad Zahir Shah (1933-1973) put forth a liberal constitution with a two-chamber legislature. This change from monarchy to democracy was a short lived reform. Unfortunately, this reform led to the growth of extremist groups reflecting ethnic, class and ideological divisions within the Afghan society [58:XVI]. The monarchy was abolished in 1973 through a military coup which established Afghanistan as a republic. Economic and social reform had little success and the republic was overthrown by a Marxist government, which ran counter to deeply rooted Afghan traditions [58:XVII]. A revolt began in 1978 igniting a countrywide insurgency. The insurgency triggered an invasion by the Soviet Union in 1979 in order to re-establish Marxist control of the government. The Soviets were unable to regain Marxist authority outside Kabul primarily due to the Afghan freedom fighters known as mujahideen. The mujahideen were trained and supplied by the United States, Saudi Arabia, Pakistan and other outside powers [58:XVIII]. The Geneva Accords in 1989 called for U.S. and Soviet noninterference in the internal affairs of Afghanistan and Pakistan, the right of refugees to return, and a full Soviet withdrawal. A note of significance is that the mujahideen were never involved in the accord agreements nor did they accept the terms. With no common enemy to fight, the ethnic, clan, religious, and personal differences within the country surfaced, spurring civil war.

The country wide turmoil gave rise to the Taliban. "Talib" means pupil, and many were educated in Pakistan madrassas. Taliban had Pashtun backgrounds and were former mujahideen. The Taliban dedicated itself to rid the country of warlords, provide order, and establish Islam within the country [58:XXI]. By 1998 the Taliban occupied 90% of the country and continued to hold an extreme, uncompromising stance on their interpretation of Islam. Given the Taliban's religious views, they committed serious atrocities against minority populations [58:XXII]. They also provided bases for other extremist groups, such as Al Queda.

Following 9/11 in 2001, the U.S., its allies, and the Northern Alliance removed the Taliban from power and began the political process for reconstruction. This operation, Operation Enduring Freedom (OEF), was a continuance of what has become the standard of life for the past 25 plus years in Afghanistan, war. 2004 saw the first presidential election, and soon after the National Assembly elections [6]. Since then a second Presidential and National Assembly election has occurred.

4.2.1 The Social Structure of Afghanistan. Due to historic and geographic factors in Afghanistan, there is a range of ethnic, linguistic, and religious diversity in the nation. While scholars have tried to classify the diversity within Afghanistan, most disagree due to different opinions of what the Afghan ethnic landscape looks like [58:103]. Often times, Afghanistan is thought of as a tribal nation. "Tribes are generally thought of as a unit of social organization that share a common ancestry and culture" [4]. However, anthropologists studying Afghanistan do not refer to tribes. Ethnicity means different things to different groups; however, every group belongs to a *quam*. A *quam* is a term of identification used to explain a myriad of affiliations, a network, families, or occupations and can mean any group of people that has something in common and acts as a single group [57:8]. Frequently, a village represents a *quam* but is not limited to a geographic setting. *Quam* can also refer to descent groups from family kin to ethnic group [58:103].

There are many organizational groups within Afghanistan to include the Pashtuns, Uzbeks, Hazaras, Tajiks, Baluchs, Kirghiz, Tatar, Nuristani as well as others. A map outlining the ethnic regions within the country is given in Figure 4.1.

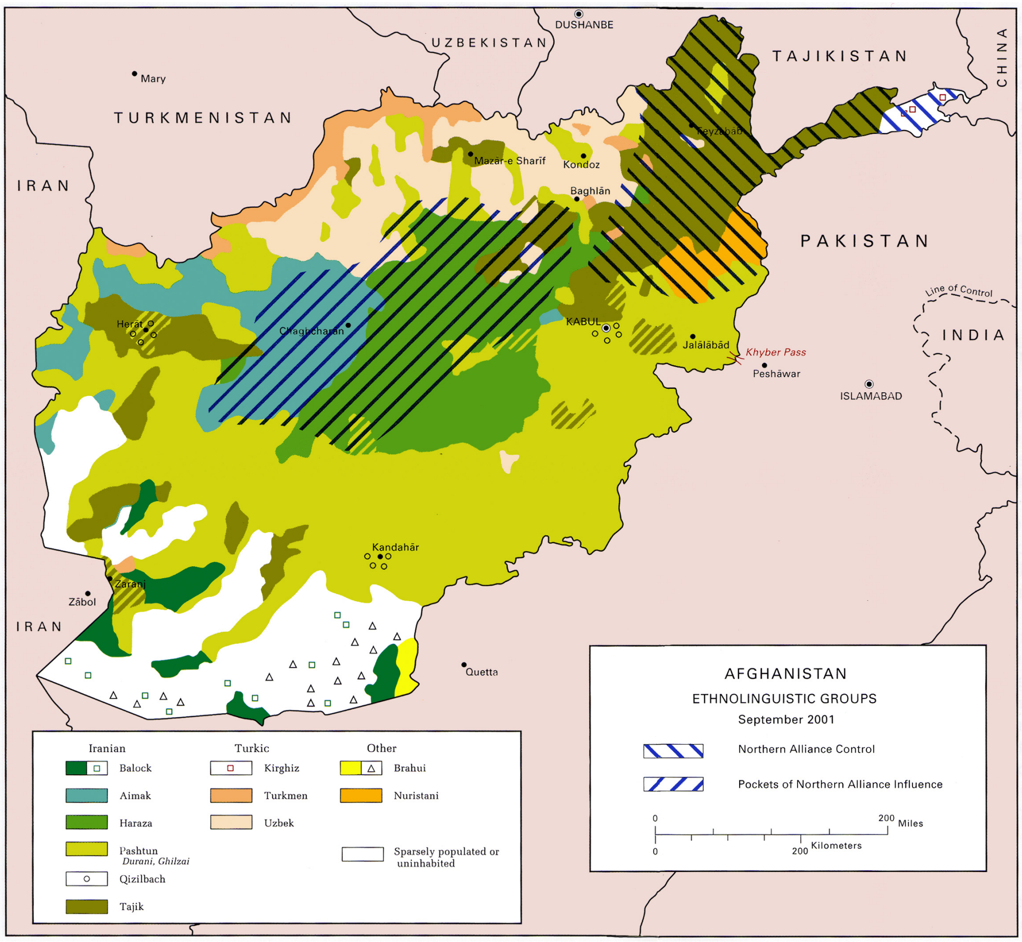


Figure 4.1 Map of Ethnolinguistic Locations in Afghanistan [52]

According to the Army's Human Terrain System (HTS), tribes in Afghanistan do not act as unified groups. They are largely non-hierarchical with no chief with whom to negotiate and are notorious for changing the form of their social organization when pressured either internally or externally [57:2]. The HTS argues that the main group acting in a tribal fashion is the Pashtun's; however, this has changed drastically over the past 30 years. The HTS asserts,

> "Pashtuns motivations for choosing how to identify and organize politically including whether or not to support the Afghan government or the insurgency are flexible and pragmatic. 'Tribe' is only one potential choice of identity among many, and not necessarily the one that guides peoples decision making" [57:2]

Additionally, there is consensus that it is hard to find groups in Afghanistan that behave in a tribal fashion according to the classic definitions of Middle Eastern tribes [57:3]. In fact, a large percentage of Afghans are not tribal at all; they (a) do not organize by kinship, (b) do have governmental institutions, and (c) do not act as one group to achieve a collective goal [57:6]. Non tribal groups include Tajiks, Uzbeks, Hazaras, and many people living within the cities.

Pashtuns behave very differently from the way tribes normally act. Pashtuns are just as likely to choose a way of organizing that has little to do with the closeness of family relations, whereas in classical tribal environments individuals would always choose to ally with blood relations [57:10]. Pashtuns will just as easily choose distant kin or non-family as allies, as they will have immediate family members. This behavior creates an unstable and unpredictable social structure as leaders will align and switch sides as time passes and events occur [57:10].

As local conflicts become part of a national conflict, many areas of Afghanistan have transformed into a more non-tribal structure. Communities were reshaped into non-family based organizations around "local strongmen, or local petty nobles", also referred to as "political entrepreneurs" [57:13]. These local strongmen form followings based on their ability to distribute resources, by charismatic leadership, or by force.

Tribal factors do play into the social structure of Pashtun society; however, they are not the foundation of social organization [57:14]. Thus, there is a need to understand the Afghanistan society from a different vantage point; SNA provides such a vantage point.

4.2.2 Afghanistan Government Structure. The constitution signed 3 January 2004 dictates the structure of the current Afghan government. The constitution serves as the legal frame work between the Afghan government and the Afghan citizens, directing for the government to unite Afghanistan. It promises to ensure that the nation belongs to all of the tribes and peoples, and pledges to honor the United Nations Charter as well as the Universal Declaration of Human Rights [1:3]. Similar to the U.S. constitution, it divides power into three branches: Executive, the National Assembly, and the Judiciary.

The elected President is responsible for implementing the tenets of the constitution [1:18], determining the policy for the country with the National Assembly, and serving as the Commander in Chief of armed forces, among other responsibilities. The government is comprised of ministers who work under the chairmanship of the president. Each minister is appointed by the president and approved by the National Assembly [1:21].

The National Assembly, as envisaged in the constitution, consists of two houses: the Wolesi Jirga (the House of the People) and the Meshrano Jirga (the House of Elders). Responsibilities of the National Assembly, found in article 90 of the Afghan constitution include [1:24]:

1. Ratification, modification or abrogation of laws or legislative decrees;
2. Approval of social, cultural, economic as well as technological development programs;
3. Approval of the state budget as well as permission to obtain or grant loans;
4. Creation, modification and or abrogation of administrative units;

5. Ratification of international treaties and agreements, or abrogation of membership of Afghanistan in them;

4.2.2.1 Afghanistan National Assembly Today. The National Assembly was formed through the first independent election in about thirty years with substantial voter turnout (53% of the country's 12.5 million registered voters - about 43% of them women) on September 18, 2005 [59]. President Karzai inaugurated the first session of the NA on December 19, 2005 and swore in the 351 members of both houses. The rough ethnic makeup of the Assembly was about 45% Pashtun, 25% Hazara, and 8% Uzbek [59]. An election for the Wolesi Jirga was held in 2010, despite delays caused by Taliban threats to candidates. According to Fazal Ahmad Manawi, the Chief of the Afghan Independent Election Commission (IEC), the finalized list of Wolesi Jirga includes 2,577 candidates with 405 women [69]. Voting occurred 18 September 2010; however results were not released by the IEC until 31 October 2010 due to accusations of fraud, vote rigging, and Taliban attacks at poll centers [41, 23]. Since almost 90% of the new Parliament members are political opponents of President Hamid Karzai [46], he tried to delay their inauguration. However, due to intense pressure from legislators and the international community, President Karzai inaugurated the new Parliament on 26 January 2011 [53].

4.2.2.2 The Wolesi Jirga (House of People). The Wolesi Jirga has 249 seats, with members directly elected by the people. Sixty-eight women were elected in 2010 to the seats reserved under the Constitution, while 17 were elected in their own rights [59]. Each province was given proportionate representation in the Wolesi Jirga according to its population. Each member of the Wolesi Jirga serves a five year term. Today, the composition of the Wolesi Jirga is 39% Pashtuns, 25% Hazara, 21% Tajik, 3% Uzbek, 3% Aymaq, 3% Arab, and the remaining is composed of Turkmen, Nuristanis, Baloch, Pashai, and Turkic [46].

4.2.2.3 The Meshrano Jirga (House of Elders). The Meshrano Jirga consists of a mixture of appointed and elected members (total 102 members). Sixty-eight members were selected by 34 directly elected Provincial Councils, and 34 were appointed by the President. President Karzai's appointments were vetted by an independent UN sponsored election board and included 17 women (50%), as required by the Constitution. Each provincial council has elected one council member to serve in the Jirga (34 members) and also in each district council (34 members). Representatives of provincial councils serve a term of four years, while representatives of district councils serve a term of three years. Sebghatulla Mojadeddi was appointed President of Meshrano Jirga. As of 20 February 2010, 68 representatives including 11 women were elected to the Meshrano Jirga [2]. This results in a total of 28 women of 102 members [2].

4.3 Data Limitations

The data collected for this thesis is described in § 3.3.2, and is entirely based upon open source research. Initial efforts were made to find governmental data from academic studies. However, limited information was found thus requiring the use of open source websites. Since the data is from open sources, gaps exist; i.e., this data does not provide a complete picture of the social structure of the current Afghanistan government. For example, many of the Afghanistan Ministries did not have websites. However, if they did, there was often limited information on the organizational structure, or who filled lead roles. Therefore, this analysis is intended as a demonstration of the technique proposed in Chapter 3, extended to government level implementation. Because of these limitations, the study should not be considered a detailed analysis of the governmental structure.

Due to the nature of research on governmental institutions in a failed or failing state, it is difficult to capture those individuals or groups that are not included within the government. In order to mitigate this issue, some businesses active in Afghanistan were included, as well as the organizations with which they are associated. For example, Kaweyan Business Development Services claims the Ministry of Women Affairs as a client, thus is tied somewhat to the government; but it may be in a structural hole to access others not directly represented in government activities. This is an example of the structural holes that exist between the government and business within Afghanistan. In view of these potential omissions, the analysis that follows should be viewed as a demonstration of how SNA and structural hole analysis might be applied. While the demonstration does reveal some interesting points, caution should be taken in applying the results of this demonstration.

4.4 Analysis

This section presents the analysis of the social network created via open source resources. The network contains 391 nodes with 462 edges. Across the network, the average degree is 2.36. This network contains nodes represented by individuals, educational institutions, organizational structure, and tribe. In order to distinguish these subnetworks, the node labeling structure is captured in Table 4.1. Each node is numerically coded from the key it was entered with into the Microsoft Access Database created for this thesis. For example, all individuals will have a node label ranging from 1 - 999.

Table 4.1 Node Label Structure

Sub Network	Node Key	# in Subgroup
Name ID	0000 - 999	152
Education ID	1000 - 1999	45
Organization ID	2000 - 2999	189
Tribe ID	3000 - 3999	5
Total		391

Figure 4.2 graphically represents all networked data. Note this graph contains nodes representing individuals, organizations, educational institutions, and tribes.

Examination of the graph indicates the presence of some highly connected actors (individuals or groups). However, as this is a representation the hierarchical structure within the government, it is not surprising. There are no disconnected subgroups in the graph because all groups in this limited open source sample are either in the government or connected to it through the Afghanistan Investment Support Agency (AISA). This agency represents an aspect of the business community that is trying to build the nation through business, separate from direct government control.

Sorting by each individual measure, the top 20 nodes are listed in Table 4.2. All data is captured by a 4 digit node label and has been converted to the node name for presentation in this study. For example, the three actors scoring the highest in degree the Provincial Government (2126), the Cabinet of Ministers (2122), and the Kaweyan Business Development Services (2210). Interestingly, it is seen that many actors who score highly across most of the measures. For example, the Provincial Government (2126), the Cabinet of Ministers (2122), Kaweyan Business Development Services (2210), Ministry of Defense (2023), Kabul University (1016), President Karzai (8), Zarar Ahmad Moqbel - Minister of Counter Narcotics (56), and Anwar ul-Haq Ahadi - Minister of Commerce and Industries (52) are actors that are commonly in the top 20 of these measures. Of particular note, the Provincial Government scores highest in six of the eight measures. This result is expected as the Provincial Government is highly interconnected.

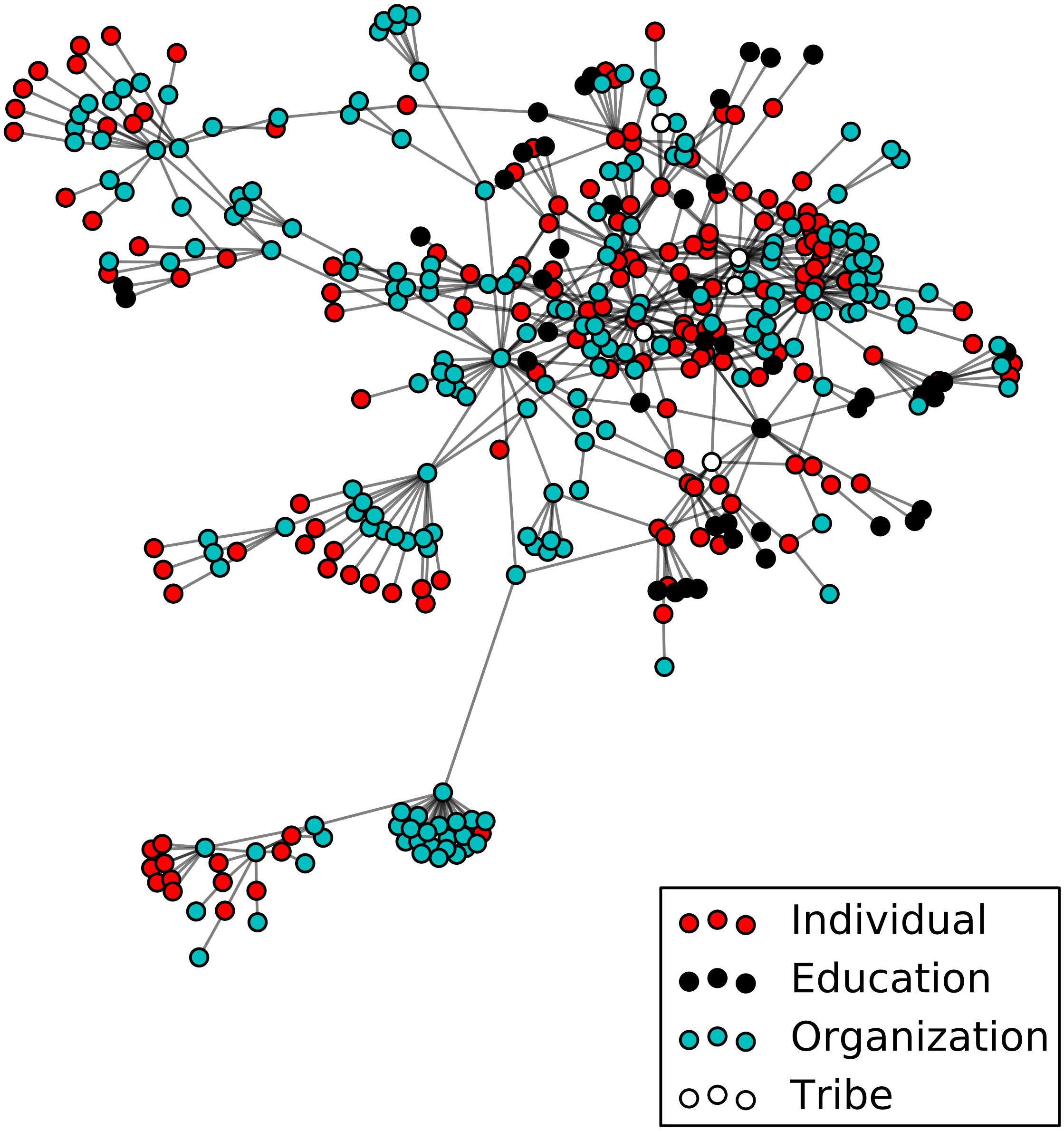


Figure 4.2 Afghanistan Data Network Layout

Table 4.2 Afghan Data Set Nodes Sorted by Measure, Top 20 Nodes

Rank	Degree	Centrality	Betweenness Centrality	Eigenvector Centrality
1	Provincial Government	Provincial Government	Cabinet of Ministers	PG
2	Cabinet of Ministers	Cabinet of Ministers	Office of President	Office of President
3	Kaweyan Business	Kaweyan Business	Provincial Government	Baghlan PG
4	Ministry of Defense	Ministry of Defense	MoWA	Balkh PG
5	DM-MFA Political Affairs	DM-MFA Political Affairs	Kaweyan Business	Bamyam PG
6	Kabul University	Pashtun	MFA	Farah PG
7	Hamid Karzai (President)	Kabul University	Ministry of Defense	Faryab PG
8	Zarar Moqbel (M-MoCN)	Mujahideen	DM-MFA Political Affairs	Kandahar PG
9	Mujahideen	Hamid Karzai (President)	Kabul University	Kunduz PG
10	Office of President	Abdul Wardak (M-MoD)	Afghan Investment Support	Nangarhar PG
11	Abdul Wardak (M-MoD)	Zarar Moqbel (M-MoCN)	MPH	Nuristan PG
12	Aid Afghanistan for Education	Office of President	Pashtun	Oruzgan PG
13	Pashtun	Aid Afghanistan for Education	Office of President	Kabul Mayor
14	Anwar Ahadi (M-MoCI)	Tajik	Istiqlal French School	Badakhshan PG
15	MPH	Karim Khalili (VP)	Abdul Wardak (M-MoD)	Badghis PG
16	Suraya Dalil (M-MPH)	MFA	Hassan Ghazanfar (M-MoWA)	Daykundi PG
17	Omar Zakhilwal (M-MoF)	Kabul Medical Institute	Hamid Karzai (President)	Ghazni PG
18	Boumi (Business)	Hezb-e Wahdat Islami	Omar Zakhilwal (M-MoF)	Ghor PG
19	Amena Afzali (M-MoLSAMD)	Ghulam Wardak (M-MoE)	Boumi (Business)	Hemland PG
20	Hassan Ghazanfar (M-MoWA)	Omar Zakhilwal (M-MoF)	Suraya Dalil (M-MPH)	Heart PG

Notes-

DM-MFA: Deputy Minister - Ministry of Foreign Affairs
DM-MoCI: Deputy Minister - Ministry of Commerce and Industries
MFA - Ministry of Foreign Affairs
M-MBTA: Minister - Ministry of Border and Tribal Affairs
M-MoCI: Minister - Ministry of Commerce and Industries
M-MoCN: Minister - Ministry of Counter Narcotics
M-MoD: Minister - Ministry of Defense
M-MoE: Minister - Ministry of Education
M-MoF: Minister - Ministry of Finance
M-MoH: Minister - Ministry of Haj (Pilgrimage)
M-MoLSAMD: Minister - Ministry of Labor, Social Affairs
M-MoWA: Minister - Ministry of Woman's Affairs
M-MPH: Minister - Ministry of Public Health
MoIA: Ministry of Islamic Affairs
MoJ - Ministry of Justice
MoM: Ministry of Mines
MoWA: Ministry of Woman's Affairs
MPH: Ministry of Public Health
MTIT: Ministry of Telecom and Information Technology
PG: Provincial Government
VP: Vice President

Table 4.2 Afghan Data Set Nodes Sorted by Measure, Top 20 Nodes (con't)

Rank	Effective Size	Efficiency	Constraint	Indirect Constraint
1	Provincial Government	Provincial Government	Provincial Government	Kamela Sidiqi (Pres-Kaweyan)
2	Cabinet of Ministers	Cabinet of Ministers	Cabinet of Ministers	Development Alternative Inc.
3	Kaweyan Business	Kaweyan Business	Kaweyan Business	GTZ (Business)
4	Ministry of Defense	Ministry of Defense	Ministry of Defense	Mercy Corps
5	DM-MFA Political Affairs	DM-MFA Political Affairs	DM-MFA Political Affairs	UN Habitat
6	Kabul University	Kabul University	Kabul University	Parwaz Microfinance
7	Hamid Karzai (President)	Office of President	Office of President	CNFA (Business)
8	Zarar Moqbel (M-MoCN)	Anwar Ahadi (M-MoCI)	Abdul Wardak (M-MoD)	Hand in Hand (Business)
9	Office of President	MPH	Hamid Karzai (President)	ASMED (Business)
10	Abdul Wardak (M-MoD)	Suraya Dalil (M-MPH)	Anwar Ahadi (M-MoCI)	Afghanaid
11	Mujahideen	Amena Afzali (M-MoLSAMD)	MPH	MTIT
12	Aid Afghanistan for Education	Hassan Ghazanfar (M-MoWA)	Suraya Dalil (M-MPH)	MoIA
13	Pashtun	Arsala Jamal (M-MBTA)	Zarar Moqbel (M-MoCN)	Inat'l Organizations Migration
14	Anwar Ahadi (M-MoCI)	Kabul Medical Institute	Pashtun	UN Development Fund for Women
15	MPH	MoM	Omar Zakhilwal (M-MoF)	Afghan Women Business Ass.
16	Suraya Dalil (M-MPH)	Mohammad Neyazi (M-MoH)	Aid Afghanistan for Education	Afghan Business Women Council
17	Omar Zakhilwal (M-MoF)	Ghulam Wardak (M-MoE)	Amena Afzali (M-MoLSAMD)	Medical
18	Boumi (Business)	DM-MoCI	Hassan Ghazanfar (M-MoWA)	Solidarity Afghan
19	Amena Afzali (M-MoLSAMD)	MFA	Arsala Jamal (M-MBTA)	International Labour Org.
20	Hassan Ghazanfar (M-MoWA)	Habibullah Ghalib (MoJ)	Kabul Medical Institute	Concern

Notes-

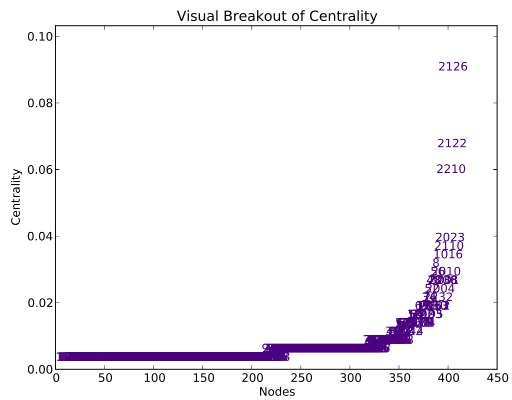
DM-MFA: Deputy Minister - Ministry of Foreign Affairs
DM-MoCI: Deputy Minister - Ministry of Commerce and Industries
MFA - Ministry of Foreign Affairs
M-MBTA: Minister - Ministry of Border and Tribal Affairs
M-MoCI: Minister - Ministry of Commerce and Industries
M-MoCN: Minister - Ministry of Counter Narcotics
M-MoD: Minister - Ministry of Defense
M-MoE: Minister - Ministry of Education
M-MoF: Minister - Ministry of Finance
M-MoH: Minister - Ministry of Haj (Pilgrimage)
M-MoLSAMD: Minister - Ministry of Labor, Social Affairs
M-MoWA: Minister - Ministry of Woman's Affairs
M-MPH: Minister - Ministry of Public Health
MoIA: Ministry of Islamic Affairs
MoJ - Ministry of Justice
MoM: Ministry of Mines
MoWA: Ministry of Woman's Affairs
MPH: Ministry of Public Health
MTIT: Ministry of Telecom and Information Technology
PG: Provincial Government
VP: Vice President

4.4.1 Ego Measure Analysis.

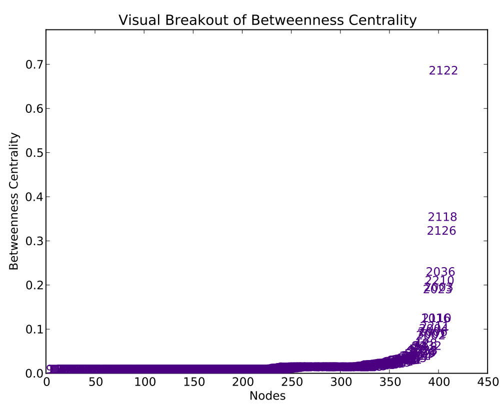
Analysis begins with the traditional measures of centrality, betweenness centrality, and eigenvector centrality. The visual breakout can be seen in Figure 4.3. This visual breakout is a plot of each node in the network along the x-axis, and that node's respective score for the given measure along the y-axis. This allows the analyst to sort and visually distinguish the separation between nodes within a given measure. It is clear from each of the graphs in this figure, there are a few actors that stand out from the rest. Centrality in Figure 4.3a breaks out actors the Cabinet of Ministers (2122), the Provincial Government (2126), and Kaweyan Business Development Services (2210). Note that these three actors have a centrality score greater than 0.05. This indicates that for a network of this size, this is a *relatively* high measure. It is to be expected that the Cabinet of Ministers, the Provincial Government, and a well connected business rate highly in centrality because they are key actors within the network.

Betweenness centrality reveals a few more actors in addition to the three that centrality broke out, and can be seen in Figure 4.3b. These actors include Ministry of Defense (2023), Ministry of Foreign Affairs (2003), Kaweyan Business Development Services (2210), Ministry of Women's Affairs (2036), Ministry of Islamic Affairs (2126), Office of the President (2118) and Cabinet of Ministers (2122). These organizations are to be expected as they are closely tied with other members of the government.

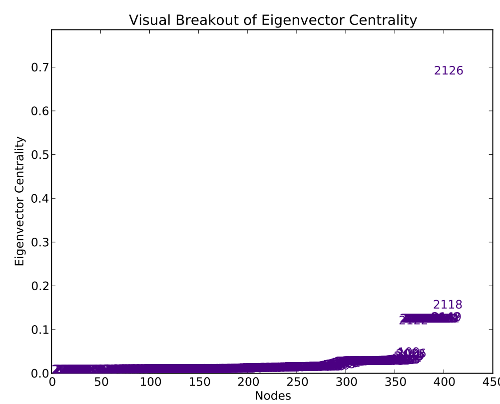
Finally, eigenvector centrality has three distinct groups within Figure 4.3c, the main body, a smaller portion above the main body, and one outlier, actor 2126 (Provincial Government). This outlier is also expected as there are many members of the Provincial Government and they are connected with other highly connected nodes within the government. A paired t-test assuming unequal variances was performed between the nodes with an eigenvector centrality of 0.10 and higher with those containing less than an eigenvector centrality score of 0.10. The test suggests that these two groups are statistically different with a p-value of 3.21×10^{-10}, or effectively zero given three significant digits. This result indicates that there are distinct separations within the dataset. The analyst can single out actors for further analysis. For the following analysis only those nodes that are statistically different in value from the remaining nodes for each ego measure are considered.



(a) Centrality



(b) Betweenness Centrality



(c) Eigenvector Centrality

Figure 4.3 Visual Breakout of Centrality, Betweenness, and Eigenvector Centrality on Afghan Data Set

Figure 4.4 represents the key actor analysis mapping betweenness centrality along the horizontal axis and eigenvector centrality along the vertical axis. The Office of the President (2118) is 1) connected to highly connected others, and 2) is close to two different groups. This implies that the Office of the President is in a position to connect disconnected groups with highly connected groups. The

Provincial Government (2126) is high in eigenvector centrality indicating this group is connected to highly connected actors. The Cabinet of Ministers (2122) is high on betweenness centrality indicating that this node is close to unconnected groups.

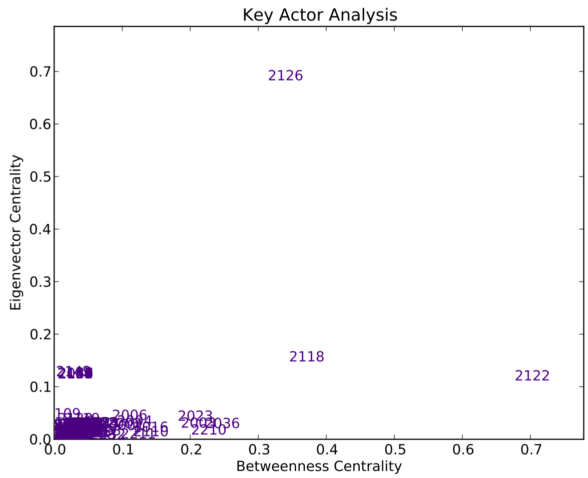


Figure 4.4 Key Actor Analysis on Afghan Data Set
2118 - Office of President
2122 - Cabinet of Ministers
2126 - Provincial Government

The key actor analysis is also represented in the color map in Figure 4.5. This graph visually depicts the node color by betweenness and the node size by eigenvector centrality. The higher the score, the darker the color and the larger the node respectively. Note that there are several large nodes illustrating how the eigenvector centrality considered more nodes to be relatively important.

From the ego measure analysis of centrality, betweenness centrality, and eigenvector centrality, it is concluded that actors Ministry of Defense (2023), Ministry of Foreign Affairs (2003), Kaweyan Business Development Services (2210), Ministry of Women's Affairs (2036), Provincial Government (2126), Office of the President (2118) and Cabinet of Ministers (2122) are relatively important within this network. Their importance is based upon how well they are connected, and to whom they are connected to (i.e., those who are important are connected to other highly connected

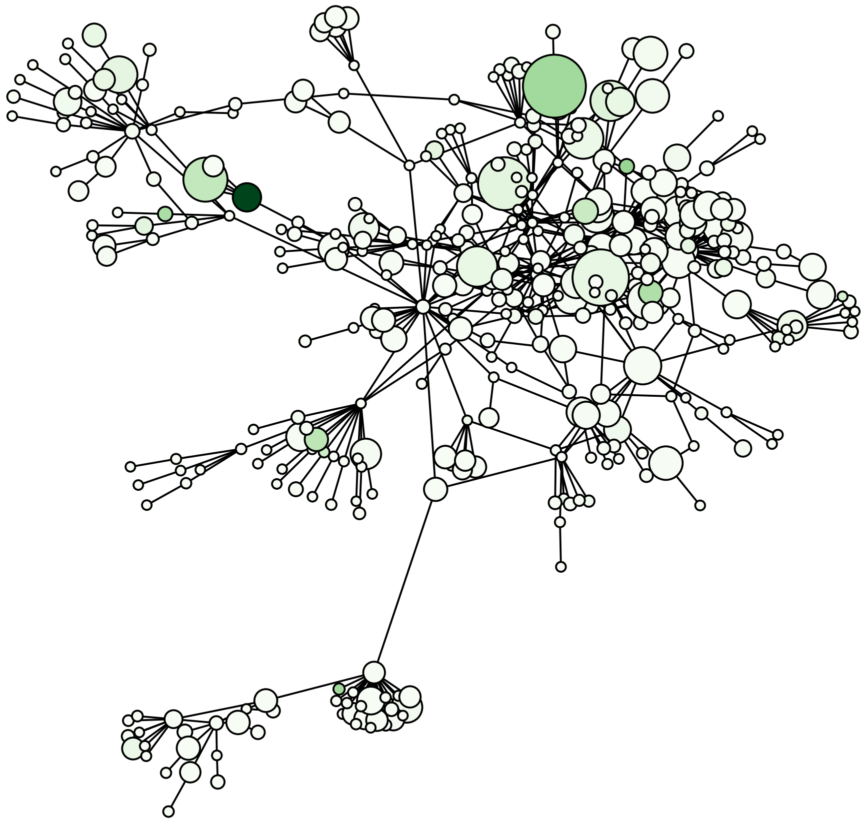
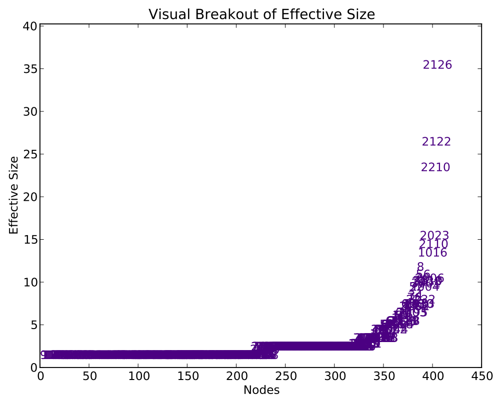


Figure 4.5 Color Map of Afghan Data Set
Node Color → Betweenness Centrality
Node Size → Eigenvector Centrality

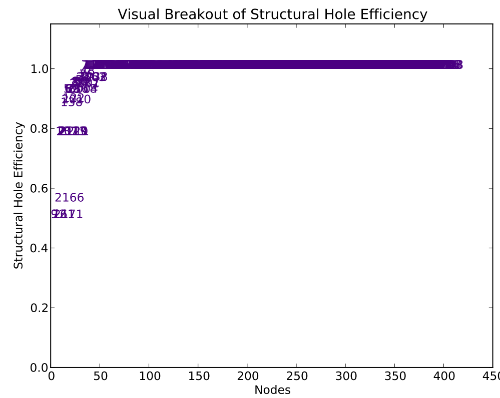
individuals). In the next section, Burt's measures of structural holes is compared to the analysis in this section. A statistical ranks test then compares all measures.

4.4.2 Structural Hole Measure Analysis. First, Figure 4.6 provides a visual breakout of Burt's four structural holes measures: effective size, efficiency, constraint, and indirect constraint.

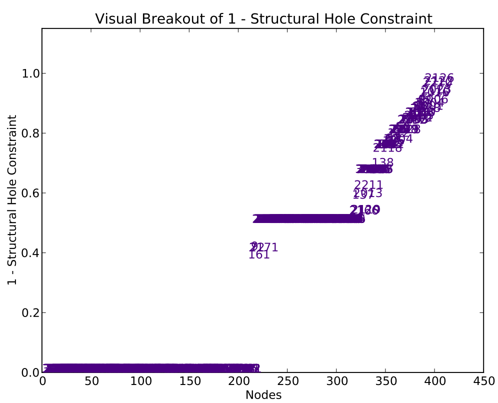
Beginning with effective size in Figure 4.6a, it is clear that it is nearly identical to the centrality measures seen in Figure 4.3a. Effective size highlights actors Provincial Government (2126), Cabinet of Ministers (2122), Kaweyan Business Development Services (2210). This indicates that each of these actors are highly connected to non-redundant contacts, creating the possibilities for spanning structural holes.



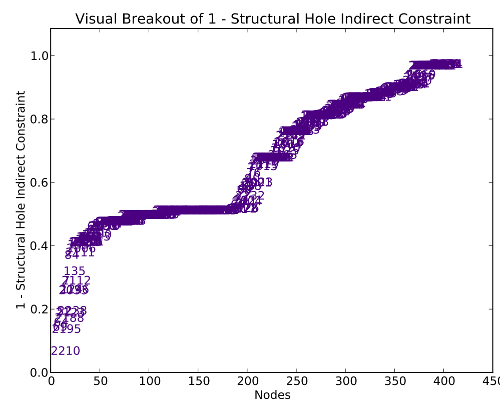
(a) Effective Size



(b) Efficiency



(c) 1− Direct Constraint



(d) 1− Indirect Constraint

Figure 4.6 Visual Breakout of Burt's Measures on Afghan Data Set

Efficiency shows a somewhat different picture in Figure 4.6b, as the majority of the actors within this data set are highly efficient. However, this result could also indicate that there are many actors with a degree of one, as any actor with only one connection is perfectly efficient because there is no redundancy. This perfect efficiency may be desirable in efficiently linking an ego's individual returns, but could indicate tenuous ties in forming a government.

Constraint is Burt's key measure. Constraint measures the redundancy that direct contacts impose upon an ego within the network. The visual breakout is seen in Figure 4.6c. Although hard to identify in the figure, the nodes with the lowest constraint in the upper right corner of the graph are the Provincial Government (2126), Cabinet of Ministers (2122), Kaweyan Business Development Services (2210), Ministry of Defense (2023), Deputy Ministry of Foreign Affairs Political Affairs (2110) and Kabul University (1016). The placement of these actors indicates that they are the least constrained by redundant relationships in the network. Moreover, many of the low constraint actors are those that betweenness centrality and eigenvector centrality marked as important actors. It is also seen that there is a large number of actors at the bottom of the figure, completely constrained by redundant relationships. While potentially undesirable in an individual's competitive network, it may be desirable when building a national government; it is desirable to be inclusive of all groups, unless the actors are from the same groups.

Finally, indirect constraint measures the redundancy that contacts of contacts impose on a given actor. The visual breakout of indirect constraint is seen in Figure 4.6d. This measure highlights actors Ministry of Public Works (2031), Ministry of Urban Development (2035), Ministry of Labor and Social Affairs (2034), Ministry of Islamic Affairs (2125) and Ministry of Telecommunication and Information Technology (2124) among others. These actors are different from those identified by previous measures as indirect constraint measures which actors are connected to actors who have low constraint. One trend of note is there are several actors with an indirect constraint score of approximately 0.50. According to Burt, this would indicate that a large number of actors are one degree away from actors who are least constrained. This signifies that they have indirect access to structural holes. Given the nature of the collection of the demonstration data set, this is not surprising.

These measures narrow the field of analysis to actors Provincial Government (2126), Cabinet of Ministers (2122), Office of the President (2118), Kaweyan Business

Development Services (2210), Ministry of Defense (2023), Deputy Minister of Foreign Affairs Political Affairs (2110) and Kabul University (1016). These actors were selected for hole signature analysis because they ranked in the top 10 of more than five of the ego measures. With this narrowed field, Burt's hole signature was performed on each actor to obtain an understanding of where SSTRO efforts could invest more resources to span structural holes and include the excluded groups of the government. The only hole signature of interest is that of the Office of the President (2118), as this actor is the only one that had varying constraining relations (i.e. the bottom line is jagged in portions), seen in Figure 4.7. The two nodes constraining Office of the President (2118) are the Legislature (2119) and the Judiciary (2120). This makes perfect sense as that is the very way the government was set up according to the constitution to balance power between the three branches. The remaining hole signature plots for the actors identified for further analysis show no constraint across relationships for the data set, (seen in Figure C.1).

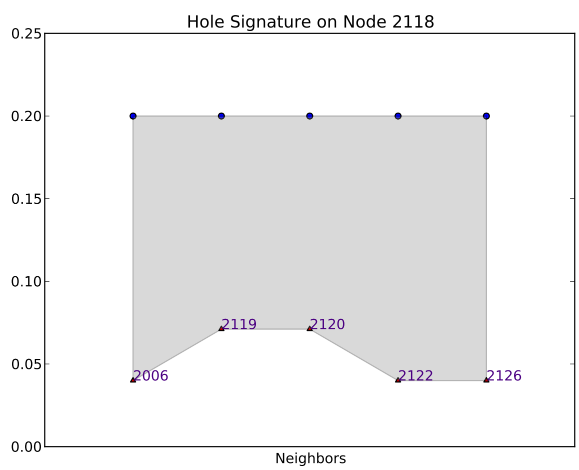
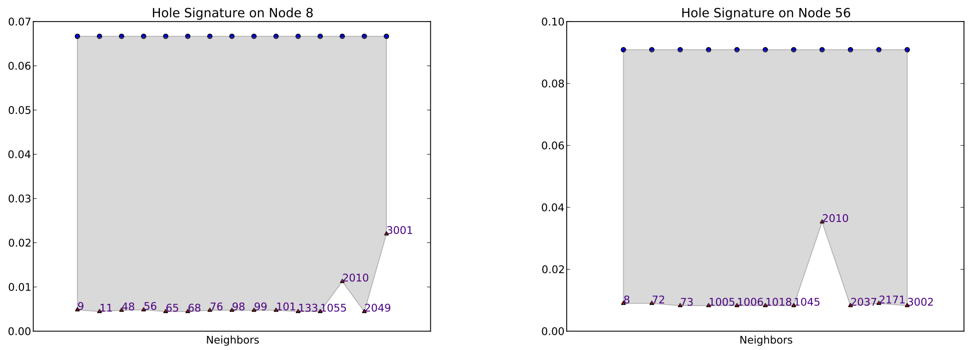


Figure 4.7 Hole Signature on Node 2118 (Office of the President)

Interestingly, Kaweyan Business Development Service (2210) is ranked in the top five in six of the eight ego measures. Kaweyan Business Development Service (2210) shows one example where structural hole analysis highlights a structural hole that can be filled in reaching many new, diverse groups for the government, thus increasing social capital. This technique ultimately can be one way to identify how to include other groups in the governmental process and increase unity within the country.

Up to this point, all the nodes for analysis have been in the 2000 series (excluding 1016 Kabul University), indicating that they are organizations. If one performs hole signature analysis on just individuals, i.e. those with node ID's between 000 - 999, then one would observe those nodes consistently ranked in the top 20 across measures. President Hamid Karzai (8), Suraya Dalil - Acting Minister of Public Health (34), Abdul Rahim Wardak - Minister of Defense (48), Anwar ul-Haq Ahadi - Minister of Commerce and Industries (52) and Zarar Ahmad Moqbel - Minister of Counter Narcotics (56) are consistent across measures excluding indirect constraint and eigenvector centrality. Indirect constraint lists only one actor as a non organization, Kamela Sidiqi, who is the president of Kaweyan Business Development Services. This is not surprising as the data did not link all business partners of Kaweyan. Eigenvector centrality placed only organizations in the top 20.

Of these five actors, two hole signatures are of interest (see Figure 4.8). Notice how both actors are constrained by the organization 2010. Organization 2010 is the mujahideen, a common bonding factor for many of those in leadership positions throughout the nation right now. During research for this study, not many of the leaders in the government had ties listed with the Taliban. From this data set, the mujahideen in the government is a very tightly knit group who are very familiar with others in the mujahideen. Mujahideen hole signature shows this group to have high redundancy. Actor 8 is the current Afghan President, Hamid Karzai and actor 56 is the current Minister of Counter Narcotics appointed by Hamid Karzai, Zarar Ahmad Moqbel. Moqbel was an active member resisting the Taliban movement and headed President Karzai's 2009 presidential re-election campaign in Parwan [55].



(a) Actor 8 (Hamid Karzai - Afghan President)

(b) Actor 56 (Zarar Ahmad Moqbel - Minister of Counter Narcotics)

Figure 4.8 Hole Signatures on Afghan Data Set for Respective Nodes

4.4.3 Statistical Testing. In order to compare across all measures used in this study, Spearman's ρ_s is calculated to determine correlation coefficients, as described in Section 2.4.1. All 21 pairwise comparisons between measures can be found in Table C.2. Table 4.3 contains a summary of the significantly correlated measures, with a p-value of less than 0.001, to three significant digits. It can be seen from this table that there is a highly negative correlation between constraint and effective size. There is also a very high correlation among betweenness, degree, and centrality. The lower yet significant correlations indicate that eigenvector centrality does not differ from degree, centrality or betweenness. This data set also confirms the findings from the example in Chapter 3; centrality, betweenness, degree, and eigenvector centrality are all highly correlated. However, this data set revealed lower correlations testing as significantly the same populations. For example, the test data set summary from Table 3.6 contained only $|\rho_s|$ values of ≥ 0.80. In Table 4.3, $|\rho_s|$ values are as low as 0.33. One possible reason for this is that the Afghan data set results in multiple ties in rank for all measures. This can lead to lower correlations, but still yield a significant result because of the normal approximation with large data sets [24:253].

Table 4.3 Significantly Correlated Spearman's ρ_s Coefficients from Table C.2

Measure	Highly Correlated with	ρ_s	p-value[a]
Constraint	Effective Size	-0.993	0.000
Constraint	Indirect Constraint	-0.469	0.000
Effective Size	Betweenness	0.336	0.000
Effective Size	Centrality	0.344	0.000
Degree	Eigenvector	0.476	0.000
Centrality	Eigenvector	0.476	0.000
Betweenness	Eigenvector	0.492	0.000
Betweenness	Degree	0.957	0.000
Betweenness	Centrality	0.966	0.000
Degree	Centrality	1.000	0.000

[a] p-values listed were $\leq 8.4 \times 10^{-12}$, or effectively 0

4.5 Applications to SSTRO

In order to apply this technique to SSTRO, there are several methods available to highlight imbalance within the government and to illustrate structural holes that exist. Mujahideen ties within the government raise the question that if there are mujahideen ties within the government, are there structural holes to fill across those ties? This section demonstrates how this technique can be utilized, given more data, to build the nation by spanning structural holes.

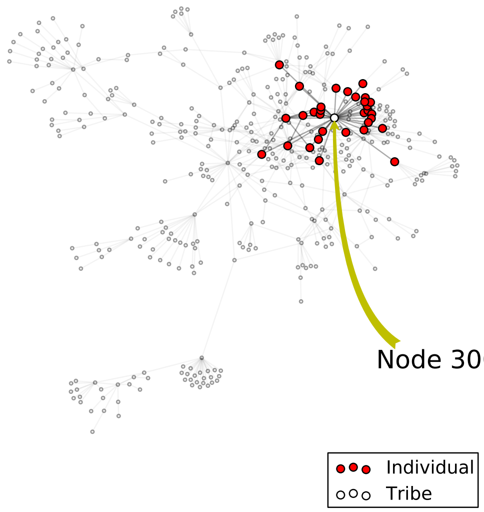
For example, focus on one of the largest degree node, the Provincial Government (2126) to investigate for structural holes that may exist within this local government. One point to note is that the provincial governors are appointed by the President, Hamid Karzai. Of interest in an ethnically fractioned society, the ethnic diversity the provincial governors are out of proportion to the national ethnic/tribal diversity as reported by the CIA world factbook [11]. This disparity can be seen in Table 4.4 where it is clear that there are more Pashtuns filling provincial governor roles than there are proportionally throughout the nation. Note that President Karzai is a Pashtun.

Table 4.4 Provincial Governors Tribe Diversity versus National Averages

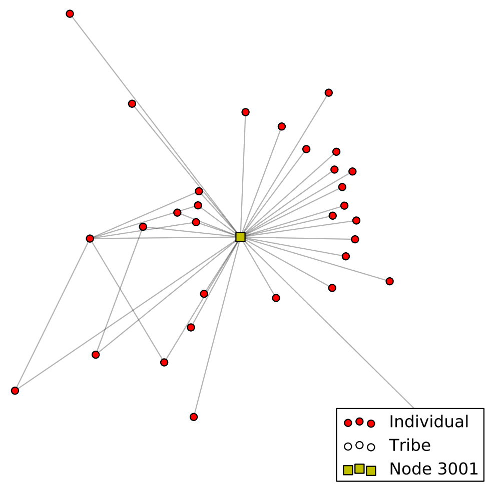
Tribe	Number	Percentage	National Percentage
Pashtun	22	65%	42%
Tajik	6	18%	27%
Hazara	5	15%	9%
Uzbek	1	3%	9%

Showing the network visually, and displaying all the Pashtuns in the government in a sub graph, (see Figure 4.9) shows the tribal disparity. On the left in Figure 4.9a is the graph of all nodes in the network, including individuals, organizations, educational insinuations, and tribes. On the right in Figure 4.9b is the Pashtun individuals in the government, highlighting their location. A note of caution must be taken when looking at this graph, as not all of the greyed out nodes are individuals, some are organizations, some are educational institutions and the proportion of Pashtun seems smaller when compared to the entire graph. Undoubtedly, the demonstration data set used in this thesis is limited and the graph of a complete data set may show more Pashtun throughout the government. Not highlighted in Figure 4.9, is a tribal connection from the Karzai family tribe of Popalzai Pashtun and the Taliban number 2, Mullah Abdul Ghani Baradar. This connection is not shown because Mullah Abdul Ghani Baradar is not linked to the government, thus he is not on the graph. This is *not* to suggest that there is a definitive link between these two men; however, it highlights the capability of the technique to identify the similarity of tribe between the head of state and the head of the Taliban.

One interesting note is that there are members of the provincial government willing to talk to the Taliban in order to communicate and create unity. For example, Mohib Ullah Samim, a Pashtun, is the Provincial Governor of Pakitka; during the ceremony for his first day in office he states,



(a) Entire Network



(b) Isolated Pashtun Network (3001)

Figure 4.9 Network of Node 3001: Pashtun Individuals Within Government

> "To the Taliban I say, come to the government and talk. Let's make it better for all of us. I will try to respect everyone. I will try to bring unity to all tribes. I will be working for unity on behalf of all people of Pakitka province and my door is open 24 hours a day if you need me." [28:Mohibullah Samim].

This outreach shows that there are individuals who are willing to fill the structural holes between the government and the Taliban in an attempt to unite the nation of fractured ideals. It is now a matter of finding those individuals willing to fill structural holes. Finding these individuals helps the Afghan nation facilitate their ability to span structural holes to build unity and prosperity within the nation. A network mapping of the government and known Taliban on which a structural hole analysis were conducted might facilitate such an identification of potential bridges, liaisons and candidates to fill structural holes.

An additional question to be asked given appropriate data is how much influence the former fighters of the mujahideen have within the government. This technique can illustrate the structural holes to be filled across former mujahideen.

Figure 4.10 shows the hole signature for the mujahideen throughout the government. With the limited data set used in this study, there were at least 15 individuals identified as former mujahideen fighters, with six of those being provincial governors (from a total of 34, or 18%), the Second Vice President to Karzai, and several prominent ministers including the ministries of Counter Narcotics and Refugees and Repatriation. Former mujahideen fighters also span into several political parties throughout Afghanistan, as seen on the lower right hand corner of Figure 4.10. These political parties include the following: Hezb-e Wahdat Islami (2163), Jamiat-e Islami (2166), Shura-e Nazar (2171), Hezbi Islami (2240), Islamic Union for the Liberation of Afghanistan (2241), as well as other pro-unification political parties. Visually, the spread of the mujahideen throughout the government and all the entities that they touch can be seen in Figure 4.11. This network of the mujahideen highlights all entities where former mujahideen are directly or indirectly (i.e. friends of friends) involved within the Afghan government. These ties may help to bridge differences. Of course, the current data does not indicate if the individuals, while all members of the mujahideen, might have belonged to rival factions.

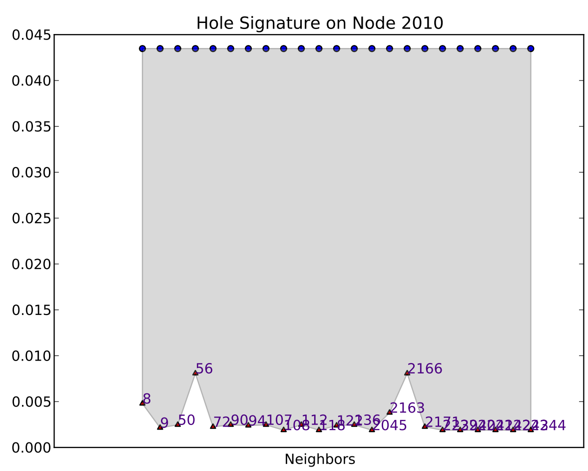
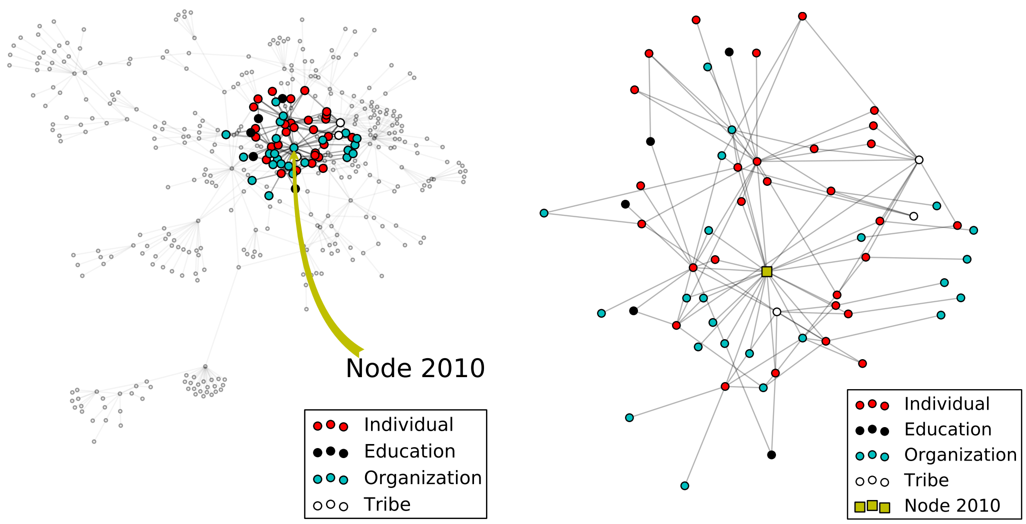


Figure 4.10 Hole Signature for 2010: Mujahideen

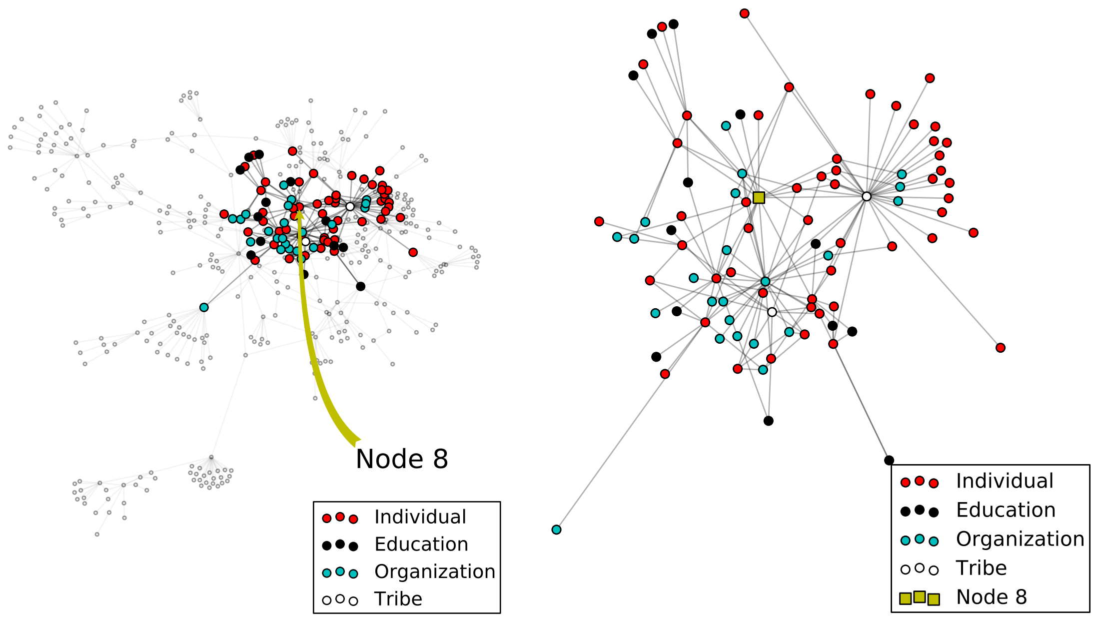


(a) Entire Network **(b)** Isolated Mujahideen Network

Figure 4.11 Network of Node 2010: Mujahideen Influence Within Government

This approach can also show how much an individual influences the rest of the government. For example, President Karzai's influence and ties to many individuals can be seen in Figure 4.12. One can see that President Karzai's influence spreads far and wide when considering all those with direct and 2^{nd} order contacts he has.

These techniques can also be used to aid isolated groups to determine how to fill these structural holes within the network. Consider a subgraph of a minority ethnicity, for example the Uzbek. It becomes apparent that there is an individual already spanning the structural hole to a majority group. In Figure 4.13, the red dot just north of node 3004 is an individual who belongs to both Uzbek and Tajik tribes. This individual filling the structural hole is Suraya Dalil - Acting Minister of Public Health (34). In fact she ranks in the top 20 in five separate measures. This illustrates that there are already individuals who fill structural holes. The impact of this can be extensive in building the nation by allowing those individuals to build ties between groups.

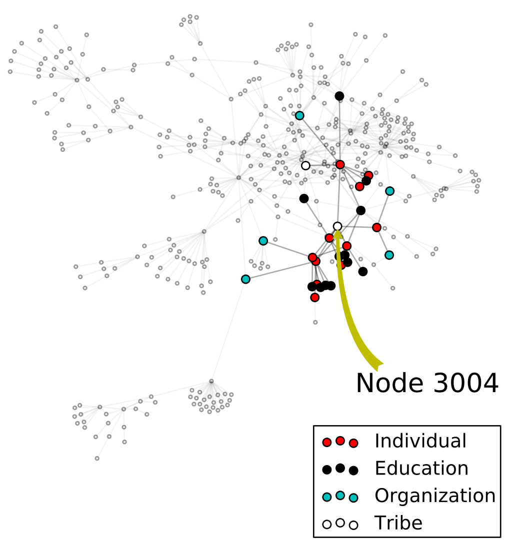


(a) Entire Network **(b)** Isolated President Karzai's Network

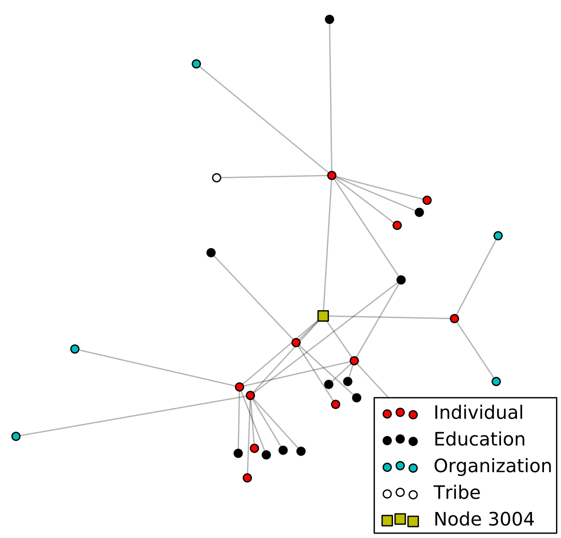
Figure 4.12 Network of Node 8: President Karzai's Influence Within Government

The techniques presented in this study assist in identifying where gaps, or structural holes, are within any social structure mapped to a social network. They can show the inclusion of the government but also highlight areas that the government may be failing to reach all peoples of the Afghan nation. With more data, such an analysis could be applied to the social structure of the nation, a province or a region. While the collection of tactical data to support operations has been the past focus of many of the behavioral studies conducted by the Coalition, turning the focus to the broader society would aid in facilitating nation building. By judiciously uncovering and filling structural holes in the fabric of society in a failed state, that state can be brought closer to a viable peace.

Overall conclusions and current applications are discussed in Chapter 5.



(a) Entire Network



(b) Isolated Uzbek Network

Figure 4.13 Network of Node 3004: Uzbek Influence

5. Conclusions and Recommendations

This chapter reviews the research objectives, discusses the immediate and long-term impact of the proof of concept, and proposes goals for follow on research.

5.1 Brief Review and Problem Statement

This study has provided a summary from the literature on factors of concern in rebuilding failed and failing states throughout the world. As evident by the 2010 FSI in Table B.4, failed and failing states are unfortunately too plentiful. There is a need for discussion and continuance of methodologies to support efforts to stabilize regions. The methodology proposed in this study is an application of SNA and a continuation of Burt's methodology of structural holes to aid in nation building efforts. By using SNA, data structures are built and relationships are codified to reveal relationships within society. In general, structural holes theory has been applied primarily to business interactions and predicting who will be promoted sooner based upon an individuals network [19]. This theory has even been applied to small, localized governments [15] to identify which projects should be funded. Prior to this study, the literature does not reveal the application of Burt's theory to an entire government, nor has it been previously suggested to use social network analysis to aid in state building.

5.2 Impact

The findings of this research improves the field of social network analysis by illustrating how it might be applied to nation building. It also provides an illustration of how the methodology might begin to be used in building a viable peace for the conflict nation. The notional analysis provided in Chapter 4 illustrates how the approach might be used to assist in building a stronger state following a conflict.

5.2.1 Immediate Impact. The utilization of the presented approach has potential for immediate impact in aiding in SSTRO. The analysis revealed that the measures effective size, centrality, betweenness centrality, degree, and eigenvector centrality were highly correlated for the data set used. These measures are statistically dependent distributions and the measures may not be entirely different from one another. This suggests the need for a more detailed analysis of these measures.

5.2.2 Long-Term Impact. The long term impact of this research is that the methodology was able to locate structural holes within the government. Revealing structural holes can aid in focusing resources to build a more inclusive state and social structure. The ultimate goal is to establish a stable secure nation through the filling of structural holes.

To effectively utilize this methodology, there is a high dependence upon an accurate and largely complete data set. The future research of social network analysts should focus on extending ways to leverage news articles and websites available through open source resources and focus whole of government efforts in collecting relevant nation building data on the target society.

5.3 Areas for Further Research

Future goals that ultimately result in support for building stable states are methods and tools to empower the individuals from within the failed and failing states to unite and establish security, economy, and the basic necessities of life. Data collection in a timely fashion is a necessary first step that will enable the social network analyst to help guide efforts in a thoughtful way in rebuilding post-conflict nations. Some future research necessary to reach this goal are outlined in this section.

5.3.1 Data Harvesting and Collection. Data collection for this research was conducted by hand which is tedious and slow. In order to expedite the process, using an automated tool to harvest data would increase the speed and efficiency of building specific data sets. While tools continue to improve, their availability and ability to focus searches that support nation building should be improved.

5.3.1.1 Database Robustness. The database created for this research should be further developed to contain a front-end input interface so the analyst will see only one screen while inputting data. In addition, existing data bases, both governmental and nongovernmental, need to be more transparent and available to all parties concerned with nation building. A critical addition would be the development of longitudinal databases on failed or failing states. Insuring the access and availability of the data would be a first step.

5.3.2 Layered Networks. Future research should try to identify ways to employ a layered network approach. This research combined all different levels of networks into one. As context does matter, future research should test the effects of identifying structural holes through layers of networks. A layer of a network can be relations, educational ties, ethnic and familial ties, organizations, common military experience and so fourth.

5.3.3 Differences of Ego Centrality Measures. One finding of this research was that effective size, centrality, betweenness centrality, degree, and eigenvector centrality were highly correlated for the data set. Future research should test the robustness of these findings to see if they apply in all situations. Determining if these measures are describing the same thing would be a benefit to the social network community by eliminating confusion, and simplifying calculations and highlight the difference in the measures.

5.3.4 Weighted Relationships. As implemented, this research removed the weights from relationships. Weights may give better insight into the amount of influence an actor carries with other actors within the network. This can bring about more robust insights as to the locations of structural holes. While it is a tenet of SNA that position in the network is what is important, other works indicate that individual characteristics, combined with structural position can have impact.

5.4 Summary

This research suggests a means of applying social network analysis techniques to nation building and SSTRO. This concept is not believed to have been applied before to nation building and SSTRO. It is a potentially ripe area for growth in order to ensure stable nations and areas. Building a viable peace in failed and failing states ultimately will secure the United States' own boarders by potential safe havens for extremist groups.

Appendix A. Generated Code

This appendix contains all code files written in Python, obtained from NetworkX website as well as code written by Capt. Bernardoni for this study. Below in Listing A.1 is the code found and modified for Structural holes, ported from Jung 2.0 and coded for use in Python to be used with Network X, written by Diedrik van Liere, RSM Erasmus University, the Netherlands. Listing A.2 contains the code used to test and validate that the code from Listing A.1 is in fact working in accordance to Burt's definitions of structural holes measures [19]. This was validated from several of Burt's examples from Figures 2.1, 2.2 [19], and Figure 2.3 [20].

Listing A.1 Structural Holes Code

```
# encoding: utf-8
"""
Functions for ego networks and structural holes.

Original code written by Diederik van Liere and Jasper Voskuilen from
RSM Erasmus University, the Netherlands and contributen to Jung 1.x
This code was ported from JUNG 2.0 (jung.sourceforge.net) and extended
by Diederik van Liere from RSM Erasmus University, the Netherlands and
the Rotman School of Management, University of Toronto.

This module contains the structural hole measures as suggested by Burt, 1992:
- Effective size
- Efficiency
- Network constraint
- Hierarchy

In addition it also calculates ego-density.

Reference: Burt, Ronald S. (1992) Structural Holes - The Social
Structure of Competition, Cambridge, MA: Harvard University Press
http://books.google.ca/books?hl=en&lr=&id=E6v0cVy8hVIC&oi=fnd&pg=PR7&dq=%22Burt%22+%22Structural+Holes%22+&ots=omMOXb-aPF&sig=fZL_7Ly4N9h805E59QbAXJatkx8#PPR9,M1
"""
#    BSD license.
__author__ = """\n""".join(['Diederik van Liere',
                             'Jasper Voskuilen',
                             'Aric Hagberg <hagberg@lanl.gov>'])
__all__= ['ego_density',
           'effective_size',
           'efficiency',
           'constraint',
           'local_constraint',
           'aggregate_constraint',
           'hierarchy'
           ]
```

```
import math
import networkx as nx

# helpers for algorithms

def all_neighbors(G,n):
    # same as neighbors for undirected graphs
    # both in- and out-neighbors for directed graphs
    if G.is_directed():
        nbrs=G.predecessors(n)+G.successors(n)
    else:
        nbrs=G.neighbors(n)
    return nbrs

def mutual_weight(G,u,v):
    try:
        w=G[u][v].get('weight',1)
    except:
        w=0
    try:
        w+=G[v][u].get('weight',1)
    except:
        pass
    return w

def normalized_mutual_weight(G,u,v,max_scaled=False):
    if max_scaled:
        mw=float(max([mutual_weight(G,u,w) for w in all_neighbors(G,u)
            ]))
    else:
        mw=float(sum([mutual_weight(G,u,w) for w in all_neighbors(G,u)
            ]))
    if mw==0:
        return 0
    return mutual_weight(G,u,v)/mw

##############


def ego_density(G,v):
    # Is there a definition that makes sense for weighted graphs?
    H=nx.ego_graph(G,v,center=False,undirected=True)
    return nx.density(H) # multiply by 100 to get percentage


def effective_size(G,n):
    """
    Burt's effective size is the number of nonredundant
    contacts for node n within the connected network G (p. 47,
    and equation 2.2 page 52).  Two contacts are redundant to
```

```
    the extent that they provide the same information
    benefits to the ego (node n).
    """
    # This treats directed graphs as undirected
    # Could be modified to handle directed graphs differently
    # Ignores weights
    ndeg=float(G.degree(n)) # number of neighbors (alters)
    # create a graph with n at center, but not included in graph
    E=nx.ego_graph(G,n,center=False,undirected=True)
    deg=E.degree() # degree of neighbors not including n
    # input by Bernardoni 10 Jan 2011 for error handling of
    # division by zero
    #try:
        # degree of n - average deg of nbrs
    return ndeg - sum(deg.values())/(ndeg)
    #except ZeroDivisionError:
#        return 0.0
#    else:
#        return result


def efficiency(G,v):
    """
    Burt's efficiency measure (Burt 1992 page 53) is effective size
    of a network divided by observed number of contacts in network,
    a number ranging from zero to one.  One indicates that every
    contact in the network is nonrdundant, while zero indicates high
    contact redundancy and therefore low efficiency.
    """
    eff = effective_size(G,v)/G.degree(v)
    return eff

def constraint(G,v):
    """
    Burt's constraint measure (equation 2.4, page 55 of Burt,
    1992). Essentially a measure of the extent to which v is invested
    in people who are invested in other of v's alters (neighbors).
    The "constraint" is characterized by a lack of primary holes
    around each neighbor.  Formally: constraint(v) = sum_{w in MP(v),
    w != v} localConstraint(v,w) where MP(v) is the subset of v's
    neighbors that are both predecessors and successors of v.
    """
    if G.is_directed():
        # Intersection of in- and out-neighbors
        nbrs=[u for u in G.successors(v) if u in G.predecessors(v)]
    else:
        nbrs=G.neighbors_iter(v)
    result = 0.0
    for n in nbrs:
        result += local_constraint(G,v,n)
    return result
```

```
def local_constraint(G, u, v):
    """
    Returns the local constraint on u from a lack of primary holes
    around its neighbor v. Based on Burt's equation 2.4. Formally:
    localConstraint(u, v) = ( p(u,v) + ( sum_{w in N(v)} p(u,w) *
    p(w, v) ) )^2 where
    N(v) = v.getNeighbors()
    p(v,w) = normalized mutual edge weight of v and w
    """
    weight = normalized_mutual_weight(G,u,v)
    r=0.0
    for w in all_neighbors(G,u):
        r += normalized_mutual_weight(G,u,w) * normalized_mutual_weight(G,w,v)
    return (weight + r)**2


def hierarchy(G,v):
    """
    Calculates the hierarchy value for a given vertex. Returns NaN when
    v's degree is 0, and 1 when v's degree is 1.
    Formally:
    hierarchy(v) = (sum_{v in N(v), w != v} s(v,w) * log(s(v,w))}) / (v.degree() * Math.log(v.degree())
    where
    N(v) = v.getNeighbors()
    s(v,w) = localConstraint(v,w) / (aggregateConstraint(v) / v.degree())
    """
    degv=G.degree(v)
    if degv==0:
        raise NetworkXError("hierarchy not defined for degree zero node %s"%v)
    v_constraint = aggregate_constraint(G,v)
    sl_constraint = 0.0
    numerator = 0.0
    for w in all_neighbors(G,v):
        sl_constraint = degv* local_constraint(G,v, w) /v_constraint
        numerator += sl_constraint * math.log(sl_constraint)
    return numerator / (degv * math.log(degv))


def aggregate_constraint(G,v,organizational_measure=None):
    """
    The aggregate constraint on v. Based on Burt's equation 2.7.
    Formally: aggregateConstraint(v) = sum_{w in N(v)}
    localConstraint(v,w) * O(w)
    """
    if organizational_measure is None:
```

```
        """
        A measure of the organization of individuals within the subgraph
        centered on v.  Burt's text suggests that this is
        in some sense a measure of how "replaceable" v is by
        some other element of this subgraph.  Should be a number in the
        closed interval [0,1].
        The default returns 1.  Users may wish to override this
        method in order to define their own behavior.
        """
        def organizational_measure(G,n):
            return 1.0
    result=0.0
    for w in all_neighbors(G,v):
        result += local_constraint(G,v,w)*organizational_measure(G, w)
    return result

def indirect_constraint(G,v):
    """
    Network around each of v's direct contacts poses some level of
    constraint and opportunity indirectly through the contact.
    Formally: indirectconstraint(v) = sum_{w in N(v)} delta(v,w)*
    constraint(w).  delta(v,w) is arithmetic average across v's contacts.
    See Burt(2010) page 300.
    """
    #delta=float(1./(len(all_neighbors(G,v))))
    result=0.0
    for w in all_neighbors(G,v):
        result += normalized_mutual_weight(G,v,w) * constraint(G,w)
    return result
```

Listing A.2 Structural Holes Analysis Code Developed by Capt. Bernardoni

```
import networkx as nx
import structural_holes as sh
import matplotlib.pyplot as plt
import numpy as np
import csv, pylab, pickle
from operator import itemgetter
from scipy import stats as s


'''
File names (all node adjacency lists):
m_name_ed -- actors, schools attended
m_name_org -- actors, organization associated with, weight
m_name_tribe -- actors, ethnicity or tribe, weight
m_relations -- actor (ego), actor (alter), weight
```

```
m_org_hierarchy -- dept, sub-dept -- a network view of traditional
    hierarchy
m_all -- unweighted combination of all sub-graphs
'''
# Set the file name to read in for analysis
#fname = 'hole_data_test_burt2010fig23'
#fname = 'hole_data_test_burtfig22'
#fname = 'hartford_drug'
#fname = 'hole_data_test'
#fname = 'm_org_hierarchy'
#fname = 'constraint_test'
fname = 'afghan_data'

# Read in file
net = nx.read_adjlist("../data/%s.txt" %fname, create_using=nx.Graph(),
     nodetype=int)
# Create a subgraph of only connected actors
subnet = nx.connected_component_subgraphs(net)[0]
# Create positioning for consistent graph layouts
f = open('./mypos.txt','r')
pos = pickle.load(f)
f.close()


def measures(net):
    '''
    Compute all calculations on SNA measures and for structural holes
    measures for graph and store in a respective dictionary.
    sorted(dict.items(),key=itemgetter(1), reverse=True)
    '''
    # Compute Burt's Effective Size Measure
    eff_size = {} #declare dictionary
    for i in net.nodes():
        eff_size.update({(i): sh.effective_size(net,i)})
    # Compute Burt's Efficiency Measure
    eff = {} #declare dictionary
    for node in net.nodes():
        eff.update({(node) : sh.efficiency(net,node)})
    # Compute Burt's Direct Constraint Measure for each node
    const={} #dict
    for c in net.nodes():
        const.update({(c) : sh.constraint(net,c)})
    # Compute Burt's Indirect Constraint Measure for each node
    ic={} #dict for indirect constraint
    for n in net.nodes():
        ic.update({n: sh.indirect_constraint(net,n)})
    # Compute eigenvector centrality
    eig=nx.eigenvector_centrality_numpy(net)
    # Compute centrality
    cent=nx.degree_centrality(net)
    # Compute betweenness centraxytext=(.7,.25)lity
```

```
    btwncent=nx.betweenness_centrality(net)
    # Compute the degree for each node
    deg = nx.degree(net)
    return eff_size, eff, const, ic, eig, cent, btwncent, deg


def srank(dict1, dict2):
    '''
    Compute spearmans rho ranking correlation coefficient between two
    given dictionaries.  Dictionaries are intended to be dictionaries
    created from measures.
    '''
    # Extract values for each measure and store in a list
    l1 = [(b) for (a,b) in dict1.items()]
    l2 = [(b) for (a,b) in dict2.items()]
    # Convert list to a numpy array
    n1 = np.array(l1)
    n2 = np.array(l2)
    # Calculate spearman's rho correlation coefficient & p-values
    rho, pval = s.spearmanr(n1, n2)
    return rho, pval


def centrality_scatter(met_dict1, met_dict2,path="",ylab="",xlab="",
    title="Key Actor Analysis",reg=False,save=False):
    """
    Function will take two dictionaries of SNA measures and plot values
        ,
    one along the x-axis and the other along the y-axis.  Reg is option
        for
    best fit line computed by min sum squares
    met_dict1 - goes along x-axis
    met_dict2 - goes along y-axis
    reg - adds a regression line (min sum squares) to the plot
    """
    # Create figure and drawing axis
    fig=plt.figure()
    ax=fig.add_subplot(111)
    # Create items so actos can be sorted properly
    met_items1=met_dict1.items()
    met_items2=met_dict2.items()
    met_items1.sort()
    met_items2.sort()
    # Grab data
    xdata = [(b) for (a,b) in met_items1]
    ydata = [(b) for (a,b) in met_items2]
    # Add each actor to plot by ID
    for p in xrange(len(met_items1)):
        ax.text(x=xdata[p], y=ydata[p], s=str(met_items1[p][0]), color=
            'indigo')
    # If adding a best fit line, use NumPy to calculate points.
```

```
    if reg:
        # Function returns y-intercept and slope.  Create function to
        # draw linear best fit line, using min squares
        slope, yint = np.polyfit(xdata, ydata,1)
        xline=plt.xticks()[0]
        yline=map(lambda x: slope*x + yint, xline)
        # Add line
        ax.plot(xline,yline, ls='--', color='grey')
    # Set new x- and y-axis limits to data
    plt.xlim((0.0, max(xdata)+(.15*max(xdata)))) # with some buffer
    plt.ylim((0.0, max(ydata)+(.15*max(ydata))))
    # Add labels
    ax.set_title(title)
    ax.set_xlabel(xlab)
    ax.set_ylabel(ylab)
    # Save figure
    if save==True:
        plt.savefig('../images/%s_cent_scatter.pdf' %fname)
    # Display figure
    plt.show()



def cluster_scatter(dict1, ylab="", xlab="Nodes", title="",
    rev=False, save=False, savn=""):
    '''
    Bernardoni version of "clustering" - a simple display of the values of the
    given dictionary sorted by size to identify if there are any common
    communities within the graph, plotted along a number line.
    rev - in order to reverse the scattering (for constraint and indirect const)
    savn - an additional string to change the save name title
    '''
    # Create list to sort items properly
    data = [(b,a) for (a,b) in dict1.items()]
    data.sort() #sort on IC values
    # Reverse order of data to display highest values first
    if rev==True:
        data.reverse()
    # Grab data
    xdata=[(b) for (a,b) in data]
    ydata=[(a) for (a,b) in data]
    x=range(len(data))
    # Create figure and drawing axis
    fig = plt.figure()
    ax= fig.add_subplot(111)
    # Place nodes on graph as text
    for p in x:
        # To display constraint measures consistant with others
        # (best value to upper right)
```

```
        if rev==True:
            ax.text(x=x[p], y=(1-ydata[p]), s=str(xdata[p]), color='indigo')
        else:
            ax.text(x=x[p], y=ydata[p], s=str(xdata[p]), color='indigo')
    # Titles for axis and graph
    ax.set_title('Visual Breakout of %s' %title)
    ax.set_xlabel(xlab)
    ax.set_ylabel(title)
    # Set new x- and y- axis limits to data
    #plt.xlim((0.0, max(xdata)+(.15*max(xdata)))) # Give a buffer
    plt.xlim(0.0, 450)
    plt.ylim((0.0, max(ydata)+(.15*max(ydata))))
    # Save figure
    if save==True:
        plt.savefig('../images/%s_cluster_plot_%s.pdf' %(fname,savn))
    plt.show()


def cluster_network(net,save=False):
    # Calculate clustering coefficients of each node (return as dict)
    clus=nx.clustering(net)
    # Get counts of nodes membership for each clustering coefficient, clean up
    unique_clus=list(np.unique(clus.values()))
    clus_counts=zip(map(lambda c: clus.values().count(c),
        unique_clus),unique_clus)
    clus_counts.sort()
    clus_counts.reverse()
    # Create a subgraph from nodes with most frequent clustering coefficient
    mode_clus_sg=nx.subgraph(net,[(a) for (a,b) in clus.items()
        if b==clus_counts[0][1]])
    # Graph the subgraph
    draw_graph(mode_clus_sg, sub='_cluster_net',save=save)


def hole_signature(net,node,save=False,legend=False):
    '''
    Create a Hole Signature graph, from Burt's definition (1992, p.66)
    Line on top is p_ij - or the allocation of time and energy
    Line on bottom is c_ij - or the constraint. This code presents the graph
    on given node.
    '''
    # Compute points to plot along x and y axis
    x = sh.all_neighbors(net,node)
    x.sort()
    xinc = range(len(x)) # Increment for x-axis to evenly spread plot
    # Porportion of time and energy invested from i to j
```

```python
    pij = [sh.normalized_mutual_weight(net,node,i) for i in x]
    # Constraint from i to j
    cij = [sh.local_constraint(net,node,j) for j in x]
    # Create figure and drawing axis
    fig = plt.figure()
    ax= fig.add_subplot(111)
    # Add each point to plot
    ax.scatter(xinc, pij, c='b', marker='o', label='$p_{ij}$ Investment')
    ax.scatter(xinc, cij, c='r', marker='^', label='$c_{ij}$ Constraint')
    # Fill in area between points
    ax.fill_between(xinc, cij, pij, facecolor='grey', alpha=0.3)
    # Add node labels to bottom points
    for p in range(len(x)):
        ax.text(x=xinc[p], y=(cij[p]), s=str(x[p]), color='indigo')
    # Housekeeping for axies and title
    ax.set_title('Hole Signature on Node %d'%node)
    ax.set_xlabel('Neighbors')
    ax.xaxis.set_major_locator(pylab.NullLocator())
    if legend==True:
        ax.legend(loc='upper right')
    # Save plot
    if save:
        plt.savefig('../images/%s_hole_sig_Node%d.pdf' %(fname,node))
    plt.show()


def draw_graph(net,sub_title='',save=False, colored=False):
    '''
    Show network on a customized figure, with colored nodes, or not.
    pos is taken from the global 'pos' calculated at program runtime
    to ensure consistent graph layout.
    '''
    # Set ranges for different types of nodes
    ind = []
    ed = []
    org= []
    tribe= []
    for i in iter(net):
        if i < 1000:
            ind.append(i)
        elif i > 1000 and i < 2000:
            ed.append(i)
        elif i > 2000 and i < 3000:
            org.append(i)
        else:
            tribe.append(i)
    # Create figure and drawing axis
    plt.figure(figsize=(8,8))
    plt.subplot(111)
```

```
    # Draw graph
    if colored==True:
        # Set color for individuals to Red
        i=nx.draw_networkx_nodes(net, pos, node_size = 30,
          nodelist=ind, node_color='r')
        # Set color for education to Black
        e=nx.draw_networkx_nodes(net, pos, node_size = 30,
          nodelist=ed, node_color='k')
        # Set color for organizations to Cyan
        o=nx.draw_networkx_nodes(net, pos, node_size = 30,
          nodelist=org, node_color='c')
        # Set color for tribes to White
        t=nx.draw_networkx_nodes(net, pos, node_size = 30,
          nodelist=tribe, node_color='w')
        # Draw edges
        nx.draw_networkx_edges(net, pos, alpha = 0.5)
        plt.sci(i)
        plt.sci(e)
        plt.sci(o)
        plt.sci(t)
        # Place legend on graph
        plt.legend( (i,e,o,t),
          ('Individual', 'Education', 'Organization', 'Tribe'),
          loc='lower right')
    # Graph with just labels
    else:
        nx.draw(net, pos, node_size=0, alpha=0.4,
          node_color="black",with_labels=True,
          font_size=12, font_color="black")
    plt.axis('off', axisbg='w')
    plt.show()
    if save:
        plt.savefig('../images/%s%s.pdf'%(fname,sub_title)) #save as pdf
        #plt.savefig('../images/%s.eps'%fname) #save as eps

def draw_ego_sub_graph(net, ego, radius, save=False):
    '''
    Routine to draw ego network scaled to entire network size,
    while maintining distance of original network.  Includes all
    neighbors of distance <= radus from ego.
    Pos is from global 'pos' in order to keep same layout.
    '''
    # Calculate subgraph of neighbors around ego
    sub_net = nx.ego_graph(net, ego, radius)
    # Create plot figure
    plt.figure(figsize=(8,8))
    plt.subplot(111)
    # Draw graph of entire network with no nodes for positioning
    nx.draw_networkx_nodes(net, pos, node_size=5,
        node_color='0.9', alpha=0.4)
```

```python
    nx.draw_networkx_edges(net, pos, alpha=0.05)
    # Set ranges for different types of nodes
    ind = []
    ed = []
    org= []
    tribe= []
    for i in iter(sub_net):
        if i < 1000:
            ind.append(i)
        elif i > 1000 and i < 2000:
            ed.append(i)
        elif i > 2000 and i < 3000:
            org.append(i)
        else:
            tribe.append(i)
    # Set color for individuals to Red
    i=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=ind, node_color='r')
    # Set color for education to Black
    e=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=ed, node_color='k')
    # Set color for organizations to Cyan
    o=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=org, node_color='c')
    # Set color for tribes to White
    t=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=tribe, node_color='w')
    # Draw nodes on graph if node exists in subgraph
    if len(ind)>0:
        plt.sci(i)
    if len(ed)>0:
        plt.sci(e)
    if len(org)>0:
        plt.sci(o)
    if len(tribe)>0:
        plt.sci(t)
    # Draw edges of nodes for ego network only
    nx.draw_networkx_edges(sub_net, pos, alpha=0.3)
    # Place pointer to ego node
    plt.annotate('Node %s'%ego, xy=pos[ego],  xycoords='data',
                 textcoords = 'figure fraction',
                 xytext=(.7,.30), #textcoords='offset points', #xytext=(.7,.25)
                 size=20,
                 #bbox=dict(boxstyle="round", fc="0.8"),
                 arrowprops=dict(arrowstyle="fancy",
                                 fc="y", ec="none",
                                 connectionstyle="angle3,angleA=0,angleB=-90"),
                 )
    # Place legend on graph
```

```python
    plt.legend( (i,e,o,t),
     ('Individual', 'Education', 'Organization', 'Tribe'),
     loc='lower right')
    plt.axis('off')
    if save:
        plt.savefig('../images/%s_%d_sub_net.pdf'%(fname,ego)) #save as pdf
    # Compute percentage of network covered by ego's network
    sp = nx.single_source_shortest_path_length(net.to_undirected(),
         ego, cutoff = radius)
    print 'Ego size: %d \n Percent Coverage: %2.2f' %(len(sp),
    float(len(sp))/float(len(net)))


def draw_ego_sub_graph_zoom(net, ego, radius, save=False):
    '''
    Routine to draw an isolated, zoomed in ego network scaled to
    entire network size, while maintining distance of original
    network.   Includes all
    neighbors of distance <= radus from ego.
    Pos is from global 'pos' in order to keep same layout.
    '''
    # Calculate subgraph of neighbors around ego
    sub_net = nx.ego_graph(net, ego, radius)
    # Create plot figure
    plt.figure(figsize=(8,8))
    plt.subplot(111)
    # Set ranges for different types of nodes
    ind = []
    ed = []
    org= []
    tribe= []
    for i in iter(sub_net):
        if i < 1000:
            ind.append(i)
        elif i > 1000 and i < 2000:
            ed.append(i)
        elif i > 2000 and i < 3000:
            org.append(i)
        else:
            tribe.append(i)
    # Set color for individuals to Red
    i=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=ind, node_color='r')
    # Set color for education to Black
    e=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=ed, node_color='k')
    # Set color for organizations to Cyan
    o=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=org, node_color='c')
    # Set color for tribes to White
```

```
    t=nx.draw_networkx_nodes(sub_net, pos, node_size = 30,
     nodelist=tribe, node_color='w')
    # Draw nodes on graph if node exists in subgraph
    if len(ind)>0:
        plt.sci(i)
    if len(ed)>0:
        plt.sci(e)
    if len(org)>0:
        plt.sci(o)
    if len(tribe)>0:
        plt.sci(t)
    # Draw large star for ego in graph
    center = nx.draw_networkx_nodes(sub_net, pos, nodelist = [ego],
                                    node_size = 100, node_color='y',
                                    node_shape= 's')
    # Draw edges of nodes for ego network only
    nx.draw_networkx_edges(sub_net, pos, alpha=0.3)

    # Place legend on graph
    plt.legend( (i,e,o,t,center),
     ('Individual', 'Education', 'Organization', 'Tribe', 'Node %d'%ego),
     loc='lower right')
    plt.axis('off')
    if save:
        plt.savefig('../images/%s_%d_zoom.pdf'%(fname,ego)) #save as pdf



def color_map(net, color_dict, size_dict, lab=False, save=False):
    """ Creates a graph with varying color and node size.
    color_dict - dictionary determining node color
    size_dict - dictionary dtermining node size
    lab - for node labels on or off
    pos - taken from global 'pos' for consisten graph layout"""
    # Adding analysis to visualization
    fig=plt.figure(figsize=(10,10))
    plt.subplot(111,axisbg='lightgrey')
    #spring_pos = nx.spring_layout(net, iterations=100)
    # Set node color intensity
    color = color_dict.items()
    color.sort()
    color = [(b) for (a,b) in color]
    # Set node size
    size = size_dict.items()
    size.sort()
    size = [((b)*2000)+20 for (a,b) in size]
    # Use matplolib's color mat for node intensity
    nx.draw(net, pos, node_color=color, cmap=plt.cm.Greens,
        node_size=size, with_labels=lab)
```

```python
    # Save image
    if save:
        plt.savefig('../images/%s_colormap.pdf'%fname)


def csv_exporter(fname):
    """
    My version of csv_exporter, clunky
    """
    # Open file to write to
    f = open('../data/%s_metrics.csv'%fname,'wb')
    wtr = csv.writer(f)
    # Create column titles for each metric
    hdr1 = ['Node', 'Degree', 'Effective Size', 'Effeciency', 'Constraint',
         'Indirect Constraint','Centrality', 'Betweenness', 'Eigenvector']
    wtr.writerow(hdr1)
    for key in net.nodes():
        aRow = [key, '%2.2f'%deg[key],'%2.2f'%eff_size[key],
              '%.2f'%eff[key], '%2.2f'%const[key],
              '%2.2f'%ic[key],'%2.2f'%cent[key], '%2.2f'%btwncent[key],
              '%2.2f'%eig[key]]
        wtr.writerow(aRow)
    f.close()

#def csv_exporter(data_dict):
#    '''
#    Takes a dictionary of measures keyed by column headers and exports
#    data as a CSV file.
#    '''
#    # Create column header list
#    col_headers=['Actor']
#    col_headers.extend(data_dict.keys())
#    # Create CSV writer and write column headers
#    writer=csv.DictWriter(open('%s.csv'%fname, 'w'),fieldnames=col_headers)
#    writer.writerow(dict((h,h) for h in col_headers))
#    # Write each row of data
#    for j in data_dict[col_headers[1]].keys():
#        # Create new dict for each row
#        row=dict.fromkeys(col_headers)
#        row["Actor"] = j
#        for k in data_dict.keys():
#            row[k]=data_dict[k][j]
#        writer.writerow(row)

#def add_metric(net,met_dict,name=''):
#    """Adds metric data to network from a dictionary keyed by node
#    labels"""
#    if(net.nodes().sort()==met_dict.keys().sort()):
```

```
#           for i in met_dict.keys():
#               net.add_node(i, {name: met_dict[i]})
#           return net
#       else:
#           raise ValueError("Node labels do not match")
```

Appendix B. Collected Data Sets

Table B.1 contains the critical values for the Spearman's ρ_s ranking correlation coefficient. It contains rejection criterion for confidence levels (α) of $0.05, 0.025, 0.01$. Below at Listing B.1 is the node-adjacency list used to test the structural holes code used with code found at Listing A.2. Table B.2 contains a list of cabinet members and their positions as of 20 January 2011. Table B.3 contains a listing of each Provincial Governor listed by province along with the date they took office as well as their ethnicity. The Failed State Index (FSI) is seen in Table B.4, listed alphabetically. This table lists the 12 indicators of a failed state. A higher number indicates a more failed state. Ranking in this table relative to other failed states, the higher the rank (e.g. 1, 2...) the more the failed the state.

Table B.1 Critical Values of Spearman's ρ_s Rank Correlation Coefficient [63]

	α Level		
n	0.05	0.025	0.01
5	0.900	-	-
6	0.829	0.886	0.943
7	0.714	0.786	0.893
8	0.643	0.738	0.833
9	0.600	0.683	0.783
10	0.564	0.648	0.745
11	0.523	0.623	0.736
12	0.497	0.591	0.703
13	0.475	0.566	0.673
14	0.457	0.545	0.646
15	0.441	0.525	0.623
16	0.425	0.507	0.601
17	0.412	0.490	0.582
18	0.399	0.476	0.564
19	0.388	0.462	0.549
20	0.377	0.450	0.534
21	0.368	0.438	0.521
22	0.359	0.428	0.508
23	0.351	0.418	0.496
24	0.343	0.409	0.485
25	0.336	0.400	0.475
26	0.329	0.392	0.465
27	0.323	0.385	0.456
28	0.317	0.377	0.448
29	0.311	0.370	0.440
30	0.305	0.360	0.432

Listing B.1 Test Data Adjacency List

```
#hole test
#adj list for each node, a touces b, c, d, etc...
1 2 19 17
2 1 3 13 14
3 4 6 10 14
4 5
6 7 8
7 9 14
8 9
9 16
10 11 12
12 20
14 15
15 16
17 18 19
```

Table B.2 Current Afghan Cabinet as of 20 January 2011 (2009-2014) [3]

Position	Name	Status
President	Hamid Karzai	Chosen by electorate
First Vice President	Mohammed Fahim	Chosen by electorate
Second Vice President	Karim Khalili	Chosen by electorate
Foreign Minister	Zalmai Rasul	Approved by Parliament
Defense Minister	Abdul Rahim Wardak	Approved by Parliament
Interior Minister	Bismillah Khan Mohammadi	Approved by Parliament
Finance Minister	Omar Zakhilwal	Approved by Parliament
Economic Minister	Abdul Hadi Arghandiwal	Approved by Parliament
Justice Minister	Habibullah Ghaleb	Approved by Parliament
Information and Cultural Affairs Minister	Sayed Makhdum Rahin	Approved by Parliament
Education Minister	Ghulam Farooq Wardak	Approved by Parliament
Higher Education Minister	Sarwar Danesh	Acting minister, not approved by Parliament
Trade and Commerce Minister	Anwar ul-Haq Ahady	Approved by Parliament
Water and Energy Minister	Ismail Khan	Acting minister, not approved by Parliament
Transportation and Aviation Minister	Mohammadulla Batash	Acting minister, not approved by Parliament
Women's Affairs Minister	Husn Banu Ghazanfar	Acting minister, not approved by Parliament
Haj and Islamic Affairs Minister	Mohammad Yousuf Neyazi	Approved by Parliament
Public Welfare Minister	Sohrab Ali Saffary	Acting minister, not approved by Parliament
Public Health Minister	Suraiya Dalil	Acting minister, not approved by Parliament
Agriculture Minister	Mohammad Asef Rahimi	Approved by Parliament
Mines Minister	Waheedullah Sharani	Approved by Parliament
Telecommunications Minister	Amirzai Sangin	Acting minister, not approved by Parliament
Rural Rehabilitation and Development Minister	Jarullah Mansoori	Approved by Parliament
Work, Social Affairs, Martyred and Disabled Minister	Amina Afzali	Approved by Parliament
Border Affairs and Tribal Affairs Minister	Arsala Jamal	Acting minister, not approved by Parliament
Urban Development Minister	Sultan Hussain	Acting minister, not approved by Parliament
Counter Narcotics Minister	Zarar Ahmad Moqbel	Approved by Parliament
Refugees and Repatriation Minister	Abdul Rahim	Acting minister, not approved by Parliament

Table B.3 Afghan Provincial Governors as of 20 January 2011

Afghan Province	Name	Took Office	Ethnicity
Badakhshan	Baz Mohammad Ahmadi	3-May-09	Tajik
Badghis	Dilbar Jan Arman Shinwari	24-Jan-09	Pashtun
Baghlan	Mohammad Akbar Barakzai	12-Jan-09	Pashtun
Balkh	Ustad Atta Mohammed Noor	Late 2004	Tajik
Bamyan	Habiba Sarabi	23-Mar-05	Hazara
Daykundi	Qurban Ali Oruzgani	15 April, 2010	Hazara
Farah	Roohul Amin	1-May-08	Pashtun
Faryab	Abdul Haq Shafaq	28-Jun-05	Hazara
Ghazni	Musa Khan	May 16 2010	Pashtun
Ghor	Sayyed Mohammad Eqbal Munib	1-Jul-05	Hazara
Helmand	Mohammad Gulab Mangal	22-Mar-08	Pashtun
Herat	Ahmad Yusuf Nuristani	1-Feb-09	Pashtun
Jowzjan	Mohammad Hashim Zare	30-Jun-05	Pashtun
Kabul	Dr. Zabihullah Mojaddidy	1-Jul-09	Pashtun
Kandahar	Tooryalai Wesa	19-Dec-08	Pashtun
Kapisa	Ghulam Qawis Abubaker	29-Jun-05	Pashtun
Khost	Abdul Jabbar Naeemi	2-Jul-05	Pashtun
Kunar	Fazlullah Wahidi	18-Nov-07	Pashtun
Kunduz	Muhammad Anwar Jegdalek	2010/2011	Tajik
Laghman	Mohammad Iqbal Azizi	18-Mar-10	Pashtun
Logar	Atiqullah Ludin	1-Sep-08	Pashtun
Nangarhar	Gul Agha Sherzai	26-Jun-05	Pashtun
Nimroz	Ghulam Dastagir Azad	3-Feb-05	Pashtun
Nuristan	Jamaluddin Badr	1-Sep-08	Pashtun
Oruzgan	Assadullah Hamdam	1-Sep-07	Pashtun
Parwan	GeneralAbdul Baseer Salangi	6-May-09	Tajik
Paktia	Juma Khan Hamdard	17-Dec-07	Pashtun
Paktika	Mohibullah Samim	15 April, 2010	Pashtun
Panjshir	Keramuddin Keram	4-Mar-10	Tajik
Samangan	Khairullah Anosh	13 April, 2010	Uzbek
Sar-e Pol	Sayed Anwar Rahmati	25 May, 2010	Hazara
Takhar	Abdul Jabbar Taqwa	16-Mar-10	Tajik
Wardak	Mohammad Halim Fidai	24-Jul-08	Pashtun
Zabul	Mohammad Ashraf Naseri	1-Mar-09	Pashtun

Table B.4 2010 Failed State Index [8]

Country	Rank	Total	Demographic Pressures	Refugees and IDPs	Group Grievance	Human Flight	Uneven Economic Development	Economic Decline	Delegitimization of the State	Public Services	Human Rights	Security Apparatus	Factionalized Elites	External Intervention
Afghanistan	6	109.3	9.5	9.2	9.7	7.2	8.2	8.3	10.0	8.9	9.2	9.7	9.4	10.0
Albania	121	67.1	5.9	2.8	4.9	7.1	5.7	6.1	6.8	5.6	5.3	5.4	6.0	5.5
Algeria	71	81.3	6.7	6.5	8.2	6.1	7.1	5.1	7.5	6.5	7.6	7.5	6.8	5.7
Angola	59	83.7	8.4	6.9	5.9	5.6	9.1	5.0	8.1	8.0	7.3	5.9	6.8	6.7
Antigua and Barbuda	127	60.9	4.7	3.4	4.5	7.3	6.1	5.5	5.3	4.6	4.7	4.6	4.0	6.2
Argentina	148	45.8	4.6	2.2	4.5	3.8	5.8	5.1	3.6	3.7	3.8	2.4	3.2	3.1
Armenia	101	74.1	5.7	6.9	6.0	7.0	6.5	5.8	6.6	5.3	6.4	5.1	7.0	5.8
Australia	168	27.3	3.5	2.5	3.4	1.2	4.2	3.2	1.5	1.8	2.0	1.4	1.5	1.1
Austria	170	27.2	2.7	2.3	3.8	1.2	4.7	2.7	1.4	1.4	1.6	1.1	1.9	2.4
Azerbaijan	55	84.4	6.2	8.1	7.9	5.7	7.3	5.9	8.0	5.5	7.2	7.3	7.9	7.4
Bahamas	132	58.9	6.2	3.2	4.7	5.8	6.4	5.0	5.5	4.4	2.8	4.8	4.8	5.3
Bahrain	133	58.8	4.5	2.6	6.5	3.5	6.0	4.0	6.7	3.1	5.4	4.7	6.1	5.7
Bangladesh	24	96.1	8.4	6.7	8.9	8.4	8.8	7.9	8.0	8.3	7.4	8.1	8.9	6.3
Barbados	135	55.4	4.0	3.2	4.9	6.5	6.7	5.4	4.1	3.1	2.8	4.5	4.5	5.7

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Belarus	82	78.7	6.7	3.7	6.4	4.8	6.7	6.7	8.7	6.2	7.9	6.2	7.8	6.9
Belgium	163	32.0	2.6	1.8	4.4	1.3	4.7	3.7	2.3	2.1	1.5	1.8	3.0	2.8
Belize	112	68.7	6.5	5.1	4.9	6.7	7.1	6.2	6.2	5.8	3.8	5.7	4.6	6.1
Benin	93	76.8	7.7	6.7	4.2	6.7	7.4	7.4	6.4	8.4	5.5	5.3	4.1	7.0
Bhutan	50	87.3	7.0	7.3	7.7	7.1	8.5	7.5	6.9	7.3	7.9	5.8	7.7	6.6
Bolivia	53	84.9	7.6	4.7	7.7	6.7	8.7	6.8	7.1	7.5	6.6	6.5	8.3	6.7
Bosnia and Herzegovina	60	83.5	5.3	7.1	8.7	5.6	7.1	5.7	8.0	5.4	5.9	7.2	9.2	8.3
Botswana	113	68.6	9.0	6.6	4.1	5.9	7.7	6.1	5.3	6.4	4.8	4.0	2.9	5.8
Brazil	119	67.4	6.3	3.7	6.2	4.8	8.8	4.0	6.2	6.0	5.4	6.7	5.1	4.2
Brunei Darussalam	117	67.6	5.4	4.2	6.6	3.8	7.8	3.7	7.7	3.5	6.9	5.9	7.4	4.7
Bulgaria	126	61.2	4.5	3.9	4.5	5.8	6.1	5.3	6.0	5.0	4.6	5.1	4.6	5.8
Burkina Faso	35	90.7	9.3	6.2	5.9	6.6	8.8	8.0	7.7	8.8	6.6	7.3	7.6	7.9
Burma	16	99.4	8.5	8.3	8.7	6.3	9.3	8.2	9.6	8.5	9.1	8.2	8.2	6.5
Burundi	23	96.7	9.4	8.4	7.8	6.5	8.4	8.2	7.6	9.0	7.7	7.1	7.9	8.7
Cambodia	40	88.7	8.0	5.3	6.9	7.9	7.1	7.7	8.7	8.3	7.7	6.4	7.7	7.0
Cameroon	26	95.4	8.2	7.6	7.5	8.1	8.7	7.0	9.0	8.0	7.8	7.8	8.7	7.0
Canada	166	27.9	3.2	2.5	3.1	2.1	4.5	2.5	1.5	1.5	1.9	1.2	2.4	1.5
Cape Verde	88	77.2	7.7	4.1	4.4	8.2	6.0	7.0	7.2	7.4	6.0	5.5	6.1	7.6
Cen. African Rep.	8	106.4	9.1	9.3	8.9	6.1	9.2	8.4	9.0	9.2	8.8	9.7	9.1	9.6
Chad	2	113.3	9.4	9.5	9.8	8.3	9.3	8.5	9.9	9.6	9.6	9.9	9.8	9.7
Chile	155	38.0	4.1	2.6	3.4	2.5	4.5	4.6	1.8	4.0	3.4	2.3	1.5	3.3
China	62	83.0	8.8	6.6	8.0	5.9	9.0	4.3	8.3	7.0	9.0	5.8	7.2	3.1

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Colombia	46	88.2	6.7	9.0	7.2	8.3	8.3	4.6	7.7	5.8	6.9	7.7	8.0	8.0
Comoros	52	85.1	7.5	3.9	5.6	6.4	6.1	7.6	8.2	8.5	6.8	7.5	8.0	9.0
Costa Rica	138	52.0	5.5	4.6	3.9	4.5	6.5	5.4	3.9	4.1	3.3	2.5	3.2	4.6
Croatia	131	59.0	4.7	5.9	5.2	4.6	5.3	6.2	4.8	3.7	4.5	4.4	4.3	5.4
Cuba	77	79.4	6.7	5.7	5.5	7.2	6.6	6.3	7.0	5.0	7.5	7.3	7.1	7.5
Cyprus	114	68.0	4.8	4.5	7.6	5.0	7.6	4.3	5.2	3.4	3.6	5.3	7.9	8.8
Czech Republic	152	41.5	3.3	2.8	3.4	4.3	4.1	4.4	3.4	3.6	3.3	2.1	3.3	3.5
Dem. Rep. of the Congo	5	109.9	9.9	9.6	8.6	8.0	9.5	8.7	8.8	9.0	9.4	9.8	8.9	9.7
Denmark	172	22.9	2.8	1.7	3.0	1.8	2.0	3.1	1.1	1.3	1.3	1.5	1.0	2.3
Djibouti	68	81.9	7.9	6.8	5.9	5.5	6.5	6.4	7.2	7.3	6.6	6.0	7.1	8.7
Dominican Republic	93	76.8	6.5	5.1	5.8	8.3	7.8	5.9	5.6	6.9	6.5	5.6	6.8	6.0
East Timor	18	98.2	8.6	9.1	7.5	6.1	7.0	8.4	9.1	8.7	7.0	8.8	8.7	9.2
Ecuador	69	81.7	6.3	6.1	6.4	7.5	8.0	6.7	7.4	7.0	5.8	6.6	7.8	6.1
Egypt	49	87.6	7.4	6.7	8.2	6.0	7.4	6.8	8.4	6.1	8.2	6.5	8.1	7.8
El Salvador	85	78.1	8.1	5.7	5.9	7.1	7.9	6.6	6.8	7.0	6.7	6.7	4.5	5.1
Equatorial Guinea	44	88.5	8.4	2.3	6.8	7.4	8.8	4.7	9.6	8.4	9.4	8.4	8.4	5.9
Eritrea	30	93.3	8.7	7.2	6.1	7.1	6.2	8.6	8.8	8.6	8.4	7.6	7.9	8.1
Estonia	140	50.7	4.5	4.2	5.0	4.1	5.2	5.0	4.5	3.3	3.3	2.6	5.5	3.5
Ethiopia	17	98.8	9.2	7.8	8.6	7.5	8.5	8.0	7.7	8.1	8.7	7.8	9.0	7.9
Fiji	74	80.5	5.9	4.2	7.4	6.6	7.5	6.7	8.9	5.5	6.7	6.8	8.2	6.1
Finland	176	19.3	2.3	1.7	1.2	2.2	1.7	3.0	0.7	1.2	1.5	1.0	1.0	1.8
France	159	34.9	3.7	3.1	5.6	1.8	5.3	3.6	1.8	1.5	2.7	1.6	2.0	2.2

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Gabon	98	75.3	7.0	5.9	3.0	6.4	7.9	5.9	7.8	6.6	6.4	5.7	7.2	5.5
Gambia	75	80.2	7.6	6.0	4.6	6.2	6.8	7.5	7.6	7.2	7.4	5.8	6.2	7.3
Georgia	37	90.4	6.2	7.8	8.4	5.8	7.2	6.5	9.0	6.4	7.3	8.0	9.1	8.7
Germany	157	35.4	3.3	4.0	4.7	2.6	4.7	3.6	2.1	1.7	2.3	2.2	2.0	2.2
Ghana	122	67.1	7.1	5.3	5.2	7.9	6.4	5.8	5.1	7.6	4.7	2.6	4.2	5.2
Greece	147.0	45.9	4.5	2.8	4.2	4.5	4.6	4.3	4.6	3.7	3.4	3.4	2.4	3.5
Grenada	123	67.0	5.8	2.9	4.2	7.6	6.7	6.1	6.4	3.9	4.6	5.4	5.8	7.6
Guatemala	72	81.2	7.4	5.6	6.8	6.7	8.0	6.9	7.1	6.8	6.9	7.2	6.3	5.5
Guinea	9	105.0	8.3	7.5	8.2	8.6	8.7	8.9	9.8	9.0	9.5	9.4	9.3	7.8
Guinea-Bissau	22	97.2	8.5	6.8	5.8	7.1	8.4	8.3	9.1	8.8	8.1	8.9	8.9	8.5
Guyana	102.0	73.0	6.1	3.6	6.2	8.0	7.7	6.9	6.8	5.3	5.2	6.6	5.1	5.5
Haiti	11	101.6	9.3	5.6	7.3	8.6	8.3	9.2	9.3	9.5	8.3	8.2	8.4	9.6
Honduras	76	80.0	7.6	4.1	5.0	6.5	8.3	7.5	7.5	6.9	6.3	7.0	6.8	6.5
Hungary	141	50.1	3.3	3.1	3.2	4.8	5.9	5.4	5.7	3.6	3.3	2.2	5.0	4.6
Iceland	165	29.8	0.8	1.1	1.0	3.0	2.3	7.2	2.0	1.5	1.9	1.1	2.0	5.9
India	79	79.2	8.1	5.2	7.8	6.5	8.7	5.1	5.8	7.2	6.1	7.6	6.2	4.9
Indonesia	61	83.1	7.2	6.5	6.3	7.3	7.9	6.7	6.9	6.7	6.5	7.3	7.1	6.7
Iran	32	92.2	6.4	8.3	8.1	7.1	7.3	5.5	9.0	5.9	9.4	8.9	9.5	6.8
Iraq	7	107.3	8.5	8.7	9.3	9.3	8.8	7.6	9.0	8.4	9.1	9.5	9.6	9.5
Ireland	173	22.4	2.0	1.6	1.0	2.0	2.8	3.3	1.6	2.4	1.5	1.4	1.5	1.3
Israel/West Bank	54	84.6	7.0	7.8	9.5	3.8	7.7	4.4	7.3	6.8	7.8	6.5	8.2	7.8
Italy	149	45.7	4.0	3.9	4.8	2.8	4.5	4.7	4.5	3.1	3.0	4.2	4.0	2.2
Ivory Coast	12	101.2	8.4	8.0	8.9	8.2	7.9	8.0	9.0	8.3	8.3	8.2	8.5	9.5

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Jamaica	119	67.4	6.0	2.8	4.5	6.4	6.5	6.8	6.8	6.2	5.5	5.8	4.0	6.1
Japan	164	31.3	4.0	1.2	3.6	2.1	2.6	3.5	1.8	1.3	3.2	2.1	2.2	3.7
Jordan	90	77.0	6.8	7.9	6.9	4.8	7.2	6.2	5.9	5.2	7.0	5.9	6.5	6.7
Kazakhstan	103	72.7	5.8	4.0	5.7	4.1	6.2	6.7	7.5	5.5	7.1	6.3	7.6	6.2
Kenya	13	100.7	9.1	8.7	8.9	7.9	8.7	7.4	9.3	8.1	8.0	7.5	8.7	8.4
Kuwait	125	61.5	5.5	4.1	5.1	4.1	6.1	3.8	6.0	3.1	6.5	4.9	7.2	5.1
Kyrgyzstan	45	88.4	7.8	5.2	7.4	7.3	7.9	7.9	8.4	6.3	7.6	7.6	7.4	7.6
Laos	40	88.7	7.9	5.9	6.8	6.7	5.8	7.3	8.3	8.1	8.7	7.4	8.5	7.3
Latvia	135	55.4	4.3	4.3	4.6	5.0	6.0	6.3	5.4	4.2	3.5	3.0	4.3	4.5
Lebanon	34	90.9	6.8	8.9	9.0	7.0	7.2	6.1	7.3	6.0	6.8	8.9	8.8	8.1
Lesotho	67	82.2	9.2	4.8	5.2	6.7	5.7	8.7	7.2	8.5	6.3	5.9	7.2	6.8
Liberia	33	91.7	8.4	8.2	6.3	6.7	8.3	8.0	7.1	8.5	6.5	6.7	8.1	8.9
Libya	111	69.1	5.7	4.3	5.8	4.2	6.9	5.3	7.3	4.2	8.3	5.2	7.1	4.8
Lithuania	146	47.8	4.3	2.9	4.0	5.0	6.0	5.7	3.9	3.2	3.3	2.2	3.2	4.1
Luxembourg	168	27.3	1.9	1.7	3.2	1.2	2.3	2.8	2.7	2.2	1.3	2.1	3.6	2.3
Macedonia	103	72.7	4.8	4.6	7.6	6.7	7.1	6.6	6.9	4.6	5.1	5.6	6.5	6.6
Madagascar	64	82.6	8.6	4.8	5.4	5.3	7.7	7.2	7.1	8.6	5.8	6.4	7.7	8.0
Malawi	28	93.6	9.2	6.5	6.2	8.4	8.3	9.2	8.1	8.6	7.3	5.4	7.8	8.6
Malaysia	110	69.2	6.3	5.0	6.6	3.9	7.0	5.1	5.9	5.0	6.8	5.9	6.3	5.4
Maldives	84	78.3	6.3	6.4	5.2	7.1	5.3	7.0	7.3	7.1	7.3	6.1	7.4	5.8
Mali	78	79.3	8.7	4.8	6.3	7.5	7.0	8.1	5.4	8.5	5.0	7.0	4.0	7.0
Malta	145	48.2	3.7	5.8	4.2	4.1	4.4	4.2	4.1	3.2	3.7	4.0	2.0	4.8
Mauritania	39	89.1	8.5	6.4	8.0	5.2	6.8	7.7	7.5	8.3	7.3	7.9	7.9	7.6
Mauritius	150	44.4	3.7	1.2	3.5	2.6	5.7	4.1	5.1	4.2	3.7	3.7	3.3	3.6

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Mexico	96	76.1	6.8	4.1	5.8	6.8	8.0	6.5	6.6	5.8	5.8	7.5	5.5	6.9
Micronesia	108	70.6	7.0	3.1	4.5	8.1	6.8	6.4	6.6	6.6	2.8	5.1	5.5	8.1
Moldova	58	83.8	6.4	4.3	6.9	7.8	6.8	7.0	7.9	6.7	6.8	7.8	8.0	7.4
Mongolia	129	60.1	5.6	1.4	4.3	2.3	5.9	5.7	6.2	5.3	6.4	4.8	5.3	6.9
Montenegro	134	57.3	4.9	4.2	6.6	2.7	4.4	4.9	4.5	3.8	5.3	4.5	5.9	5.6
Morocco	90	77.0	6.8	6.6	6.6	6.4	7.6	6.5	7.2	6.6	6.8	5.4	6.2	4.3
Mozambique	69	81.7	8.8	3.5	4.8	7.8	7.5	7.8	7.5	8.9	7.3	6.2	5.4	6.2
Namibia	100	74.5	7.5	5.7	5.6	7.5	8.9	6.5	4.8	6.9	5.8	5.6	3.7	6.0
Nepal	26	95.4	8.1	7.0	9.2	6.2	9.0	8.3	8.1	7.6	8.7	7.7	8.5	7.0
Netherlands	166	27.9	2.7	3.2	4.7	1.9	3.2	3.0	1.2	1.5	1.3	1.1	1.7	2.4
New Zealand	171	23.9	1.5	1.4	3.3	2.1	4.3	4.0	1.0	1.6	1.5	1.1	1.2	0.9
Nicaragua	65	82.5	6.8	5.0	6.3	6.9	7.9	7.9	7.6	7.6	6.2	6.5	7.0	6.8
Niger	19	97.8	9.6	6.5	8.0	6.5	7.8	9.2	8.9	9.7	8.5	7.3	7.6	8.2
Nigeria	14	100.2	8.4	5.8	9.5	8.1	9.3	6.9	9.4	9.1	8.8	9.3	9.4	6.2
North Korea	19	97.8	8.5	5.6	7.2	5.0	8.8	9.6	9.9	9.6	9.5	8.1	7.8	8.2
Norway	177	18.7	1.7	1.6	1.3	1.2	2.4	2.6	0.8	1.1	1.6	1.2	1.1	2.1
Oman	144	48.7	4.7	1.1	3.0	1.7	2.7	4.5	6.0	4.5	6.7	5.2	6.6	2.0
Pakistan	10	102.5	8.1	8.9	9.4	7.9	8.4	6.2	8.9	7.3	8.9	9.7	9.5	9.3
Panama	130	59.3	6.3	3.5	4.4	5.0	7.5	5.6	4.8	5.5	4.5	5.2	3.0	4.0
Papua New Guinea	56	83.9	7.5	4.2	7.1	7.7	9.0	6.3	7.8	8.3	6.3	6.5	7.1	6.1
Paraguay	106	72.1	6.2	1.5	6.3	5.8	8.0	6.2	8.3	5.8	6.7	5.9	7.5	3.9
Peru	92	76.9	6.4	4.5	6.7	7.0	8.0	5.6	6.9	6.5	5.5	7.4	6.9	5.5
Philippines	51	87.1	7.7	6.7	7.6	7.0	7.4	5.8	8.6	6.3	7.5	7.9	8.0	6.6
Poland	142	49.0	4.7	3.2	3.3	5.9	4.8	5.0	4.5	3.7	3.8	2.4	3.7	4.0

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Portugal	162	33.1	3.7	1.8	2.6	2.2	3.7	4.7	1.9	3.6	3.5	1.4	1.2	2.8
Qatar	139	51.8	4.5	3.0	5.2	3.4	5.3	4.1	6.3	2.6	4.7	2.7	5.0	5.0
Rep. of the Congo	31	92.5	8.7	7.7	6.3	6.4	8.1	7.8	9.1	8.6	7.7	7.6	7.1	7.4
Romania	128	60.2	5.4	3.2	5.6	4.9	5.6	5.6	6.0	4.8	4.3	4.1	5.2	5.5
Russia	80	79.0	6.7	5.4	7.1	6.0	7.9	5.1	8.1	5.5	8.0	6.8	7.6	4.8
Rwanda	40	88.7	9.1	7.0	8.5	7.0	7.2	7.0	7.5	7.4	7.5	5.0	8.0	7.5
Samoa	107	71.1	6.9	3.1	5.1	8.0	6.6	6.2	6.4	5.1	4.5	5.8	5.3	8.1
Sao Tome	97	75.8	7.5	4.1	5.1	7.0	5.9	7.3	7.3	7.3	5.1	6.0	6.7	6.5
Saudi Arabia	87	77.5	6.3	6.2	7.8	3.5	7.3	3.1	8.2	4.1	9.1	7.8	7.8	6.3
Senegal	99	74.6	7.6	6.2	6.1	5.8	7.0	6.2	5.9	7.4	6.0	6.3	4.2	5.9
Serbia/Kosovo	86	77.8	5.6	6.9	7.8	5.3	6.9	6.2	6.8	5.2	5.6	6.5	8.0	7.0
Seychelles	115	67.9	6.1	4.3	5.0	4.5	6.9	5.8	7.0	4.5	5.9	5.6	6.0	6.3
Sierra Leone	28	93.6	9.1	7.1	6.7	8.3	8.8	8.6	7.7	9.1	6.8	5.9	7.8	7.7
Singapore	160	34.8	2.8	0.9	2.9	2.5	3.1	3.7	4.2	1.7	4.4	1.5	4.1	3.0
Slovakia	143	48.8	4.1	2.2	4.8	5.2	5.6	5.0	4.1	3.8	3.8	2.1	3.9	4.2
Slovenia	156	36.0	3.4	1.4	3.4	3.3	5.0	4.0	2.8	3.0	3.0	2.8	1.3	2.6
Solomon Islands	43	88.6	8.3	4.8	7.0	5.4	7.9	8.0	8.1	8.2	6.8	7.0	8.0	9.1
Somalia	1	114.3	9.6	10.0	9.7	8.3	8.0	9.6	10.0	9.6	9.9	10.0	10.0	9.6
South Africa	115	67.9	8.4	7.0	5.6	4.4	8.5	5.0	5.8	5.5	4.7	4.1	5.9	3.0
South Korea	153	41.3	3.6	3.3	3.9	4.8	2.5	2.8	3.9	2.3	2.8	1.5	3.6	6.3
Spain	151	43.5	3.7	2.8	6.3	1.8	5.0	4.4	1.6	2.4	2.5	5.3	5.7	2.0
Sri Lanka	25	95.7	7.3	9.4	9.6	6.7	8.7	5.9	8.6	6.4	8.8	8.5	9.4	6.4
Sudan	3	111.8	8.8	9.8	9.9	8.7	9.5	6.7	9.9	9.3	9.9	9.8	9.9	9.6

Continued on next page...

Country	Rank	Total	DP	R	GG	HF	UED	ED	D	PS	HR	SA	FE	EI
Suriname	105	72.5	6.0	3.7	6.4	6.7	7.7	6.6	6.5	5.1	5.8	6.0	5.8	6.2
Swaziland	63	82.8	9.1	4.2	4.2	6.2	6.2	8.2	8.6	7.6	7.7	6.6	6.9	7.3
Sweden	175	20.9	2.7	2.7	1.3	1.8	2.1	2.2	0.8	1.3	1.8	1.3	1.3	1.6
Switzerland	174	21.8	2.4	1.5	3.3	1.8	2.6	2.4	1.0	1.4	2.2	1.2	1.0	1.0
Syria	48	87.9	5.9	8.9	8.3	6.6	7.8	6.3	8.6	5.5	8.8	7.6	7.8	5.8
Tajikistan	38	89.2	8.0	6.2	6.9	6.3	7.1	7.5	8.9	7.3	8.7	7.3	8.4	6.6
Tanzania	72	81.2	8.2	7.3	6.4	6.1	6.7	7.2	6.5	8.3	5.9	5.6	6.0	7.0
Thailand	81	78.8	6.7	6.7	7.8	4.7	7.5	4.3	8.0	5.4	7.0	7.4	8.0	5.3
Togo	47	88.1	8.0	6.2	5.6	7.0	7.6	8.0	7.5	8.4	7.7	7.6	7.6	6.9
Trinidad and Tobago	124	66.1	5.6	3.1	4.9	7.3	7.2	4.8	5.9	5.2	5.4	6.0	5.6	5.1
Tunisia	118	67.5	5.7	3.4	5.4	5.2	7.0	5.0	6.4	5.7	7.5	6.5	6.0	3.7
Turkey	89	77.1	6.3	6.3	8.0	4.8	7.8	5.8	6.0	5.4	5.5	7.4	7.8	6.0
Turkmenistan	65	82.5	6.8	4.6	6.3	5.4	7.4	6.6	8.4	7.0	9.0	7.7	7.7	5.6
Uganda	21	97.5	8.7	8.9	8.5	6.9	8.4	7.2	7.9	8.2	7.6	8.7	8.6	7.9
Ukraine	109	69.5	5.6	3.1	6.9	6.6	6.2	6.3	7.2	4.0	5.3	3.8	7.9	6.6
United Arab Emirates	137	52.4	4.4	3.2	4.7	3.3	5.7	3.9	6.7	3.4	5.9	2.7	4.0	4.5
United Kingdom	161	33.9	3.2	3.0	4.1	1.8	4.5	3.0	1.6	2.3	2.3	2.7	3.2	2.2
United States	158	35.3	3.1	3.2	3.4	1.1	5.4	4.0	2.5	2.5	3.7	1.6	3.3	1.5
Uruguay	153	41.3	4.3	1.3	2.0	5.6	5.0	4.0	2.6	3.4	2.5	3.4	3.0	4.2
Uzbekistan	36	90.5	7.7	5.1	7.4	6.6	8.5	7.0	8.5	6.4	9.3	8.8	9.0	6.2
Venezuela	82	78.7	6.3	5.1	6.8	6.7	7.6	5.8	7.2	6.1	7.2	6.7	7.5	5.7
Vietnam	95	76.6	6.9	5.2	5.3	5.9	6.5	6.6	7.3	6.4	7.3	6.0	7.0	6.2

Continued on next page...

Yemen	15	100.0	8.6	8.3	8.2	7.2	8.6	7.9	8.7	8.6	8.0	8.9	9.2	7.8
Zambia	56	83.9	9.0	7.3	5.4	7.1	7.3	8.0	7.5	8.0	5.9	5.0	6.1	7.3
Zimbabwe	4	110.2	9.4	8.6	8.8	9.7	9.4	9.6	9.6	9.4	9.5	9.2	9.5	7.5

Appendix C. Results

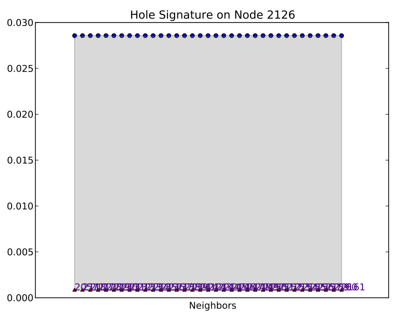
This appendix contains tables and data too long to include within the main document. Table C.1 contains the Spearman correlation coefficient for the test data set explained in Chapter 3. Figure C.1 contains additional hole signature charts from Chapter 4. These hole signatures are from some of the top organizations when all the measures are sorted by rank and do not reveal any additional information and are included for completeness. Table C.2 contains the Spearman correlation coefficient for the Afghan data set analyzed in Chapter 4.

Table C.1 Spearman ρ_s Correlation Coefficient Pairwise Comparisons - Test Data Set

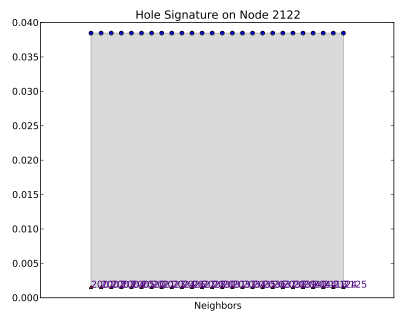
	Effective Size		Effeciency		Constraint		Indirect Constraint		Centrality		Betweenness		Eigenvector	
	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value
Degree	0.968	2.60E-12	-0.535	0.151	-0.889	1.58E-07	0.047	0.845	1.000	0.00	0.888	1.70E-07	0.884	2.32E-07
Effective Size	-	-	-0.327	0.159	-0.939	8.87E-10	-0.049	0.837	0.968	2.61E-12	0.895	9.91E-08	0.909	2.94E-08
Effeciency			-	-	0.176	0.458	-0.307	0.188	-0.535	0.015	-0.393	0.086	-0.295	0.207
Constraint					-	-	0.133	0.577	-0.889	1.58E-07	-0.815	1.20E-05	-0.892	1.27E-07
Indirect Constraint							-	-	0.047	0.845	0.223	0.345	-0.265	0.259
Centrality									-	-	0.888	1.70E-07	0.884	2.32E-07
Betweenness											-	-	0.823	8.18E-06

Figure C.1 Hole Signature on Selected Nodes from Afghan Data Set

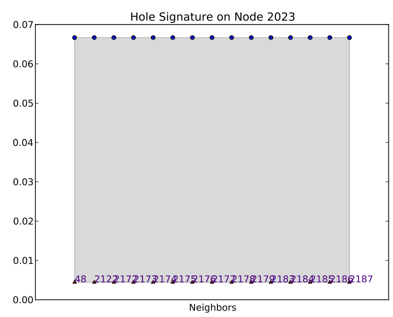
(a) Node 2126



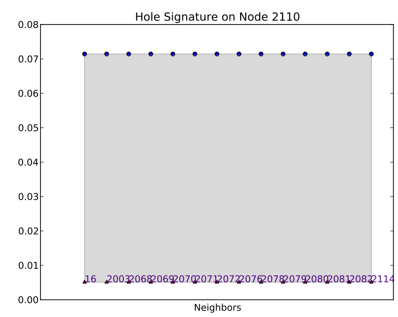
(b) Node 2122



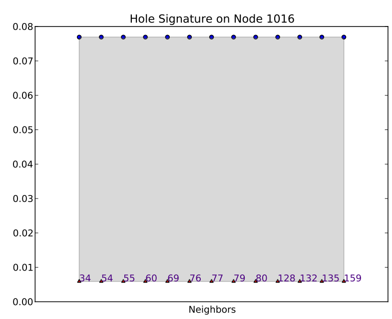
(c) Node 2023



(d) Node 2110



(e) Node 1016



(f) Node 2210

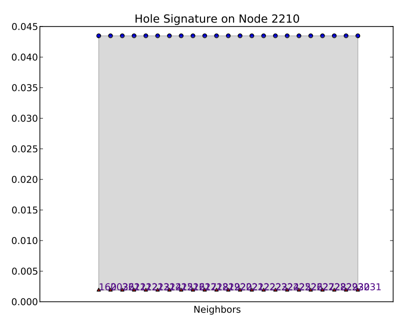


Table C.2 Spearman ρ_s Correlation Coefficient Pairwise Comparisons - Afghan Data Set

	Effective Size		Effeciency		Constraint		Indirect Constraint		Centrality		Betweenness		Eigenvector	
	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value	ρ_s	p-value
Degree	0.344	2.75E-12	-0.144	0.004	-0.345	2.14E-12	0.173	0.001	1.000	0.000	0.957	0.000	0.476	1.61E-23
Effective Size	-	-	-0.343	2.96E-12	-0.993	0.000	0.478	6.35E-24	0.344	2.75E-12	0.336	8.40E-12	0.113	0.026
Effeciency			-	-	0.389	1.53E-15	0.061	0.232	-0.144	0.004	-0.105	0.038	-0.050	0.322
Constraint					-	-	-0.469	9.65E-23	-0.345	2.14E-12	-0.334	1.20E-11	-0.113	0.025
Indirect Constraint							-	-	0.173	0.001	0.187	0.000	0.071	0.158
Centrality									-	-	0.966	0.000	0.476	1.61E-23
Betweenness											-	-	0.492	3.32E-25

Appendix D. Blue Dart

The blue dart for this study follows.

Utilizing Social Network Analysis in Support of Nation Building

— BLUE DART —

Brandon J. Bernardoni[1], Captain, USAF

March 2011

One of the greatest threat to our national security comes from fragile states unable or unwilling to provide for the needs of the people. Aiding post conflict nations building capacities, as well as assisting failing states, helps to ensure United States national security. The United States requires a comprehensive approach to foster development in these nations in ways that eliminate or at least mitigates the requirement for future military intervention. To aid in this comprehensive, whole of government approach, models and tools are needed to assist planners, commanders and decision makers.

Social network analysis provides techniques and tools utilized by social scientists to study the formal and informal interrelations in a community. Since 9/11 these techniques and tools have been increasingly utilized by the defense and intelligence communities to analyze terrorist networks to aid in thwarting foes. This study investigated the use of social networks, particularly the concept of structural hole theory, to facilitate nation building in failed and failing states. Structural holes, are gaps in the connections present in a social network. Initially developed to analyze an

[1]Masters Student, Department of Operational Sciences, Air Force Institute of Technology, Dayton, OH

individual's or a corporations formal and informal structure to increase competitive position, this study extends the concepts in structural hole theory to aid in modeling the fractured society within a failed or failing state. Investigating the social structure of a community identifies gaps between government capacity and public needs. With such knowledge, Security Stabilization Transition and Reconstruction Operations (SSTRO) can be focused to strengthen the host nation government or other elements of society. A strong, unified government provides security and unity for its citizens, especially from insurgent forces.

Identifying individuals in the professional and governmental community highlights which individuals control power within the nation, both formally and informally. Further, this technique highlights gaps within the government and illustrates how to bridge gaps between marginalized sub-groups and the central elements of a society and its government.

Applying these techniques to a national level aids commanders and international aid to focus resources and energy effectively to aid in building a stable nation state. This technique can highlight national, regional, or local gaps that can be filled to facilitate nation building, and ultimately aid in the security of the nation.

Appendix E. Quad Chart

The storyboard for this study is found below.

UTILIZING SOCIAL NETWORK ANALYSIS IN SUPPORT OF NATION BUILDING

Capt Brandon Bernardoni
Advisor: Dr. Richard F. Deckro
Department of Operational Sciences (ENS)
Air Force Institute of Technology

Introduction

Social network analysis (SNA) is a powerful set of techniques used by social scientists to study the formal and informal interrelations in a community. Since 9/11 these techniques have been increasingly utilized by the defense and intelligence communities to analyze terrorist networks to aid in thwarting foes. This study investigates the use of social networks and structural hole theory to facilitate nation building in failed and failing states. By investigation of the underlying social structure of a community, identifying structural holes and gaps within the government or other elements of the society, Security Stabilization Transition and Reconstruction Operations (SSTRO) efforts can be focused to strengthen the host nation to provide security and stability for its citizens.

This investigation focused on exploring techniques that link individuals in the professional and governmental community. Applying methodology to a national level in order to identify structural gaps within an ethnically fractured failing state. This technique can highlight national, regional, or local holes that can be filled to facilitate nation building.

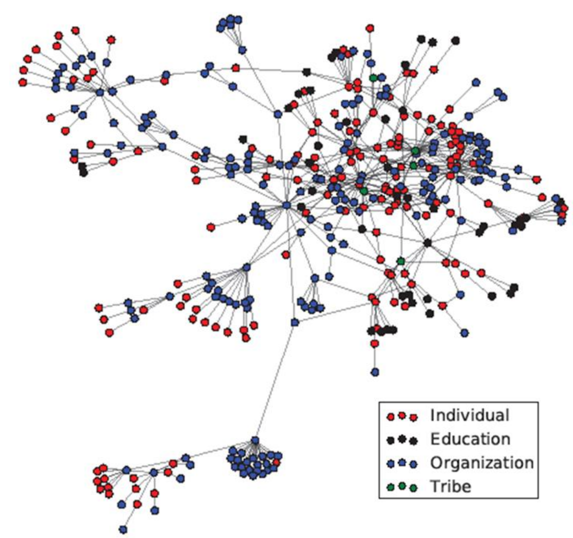
Research Goals

- Apply known techniques from SNA to support efforts to aid building failed and failing states.
- Identify gaps within the governmental or social structure of a failed or failing state
- Demonstrate SNA methods when applied to a strategic, national level

AIR FORCE INSTITUTE of TECHNOLOGY

General Framework

Social Network of Government

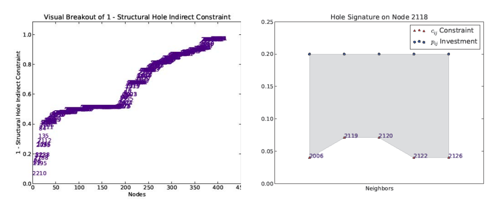


- •391 Nodes
- •462 Relationships
- •210 Nodes with degree of 1

Measures Employed

- Degree Centrality
- Eigenvector Centrality
- Betweenness Centrality
- Effective Size
- Structural Efficiency
- Structural Constraint

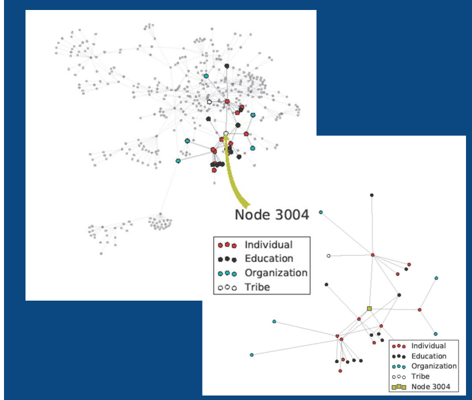
Identify Structural Holes



Application

Build stronger nations to reduce insurgent safe havens

Include excluded subgroups



Motivation

- Need for models to understand and support nation building efforts within failed states
- Provide methodology to understand how fractured subgroups interact

Impacts/Contributions

- Illustrates methodology to assist building stronger post conflict nations
- Demonstrates how to identify and understand underlying social interaction
- Aid in focusing resources to build inclusive state and social structure

Bibliography

1. The constitution of afghanistan. http://www.supremecourt.gov.af/PDFiles/constitution2004_english.pdf, 3 January 2004. accessed 17 October 2010.
2. Afghanistan Meshrano Jirga historical archive of Parliamentary election results. http://www.ipu.org/parline/reports/2382_E.htm, 2008. Accessed 28 February 2011.
3. The Embassy of Afghanistan in Washington D.C. http://www.embassyofafghanistan.org/ministers.html, December 2009. Accessed 20 January 2011.
4. Complex operations wiki. http://complexoperations.org/cowiki/Main_Page, October 2010. Joint venture by Human Social Cultural Behavior Modeling and Department of Defense. Accessed on 22 January 2011.
5. Corruption perceptions index. Technical report, Transparency International, October 2010.
6. Facts about afghanistan. http://worldfacts.us/Afghanistan.htm, Nov 2010.
7. Python programming language. www.python.org, 2010. accessed 11 Jan 2011.
8. US fund for peace: Promotion sustainable security. http://www.fundforpeace.org/web/index.php?option=com_content&task=view&id=99&Itemid=140, 2010. Accessed 6 February 2011.
9. Foreign policy: The failed states index 2010. http://www.foreignpolicy.com/failedstates, 2011. Accessed 6 Feb 2011.
10. SciPy statistics: Scientific tools for python. http://www.scipy.org/SciPyPackages/Stats, 13 February 2011. Accessed on 13 February.
11. The Central Intelligence Agency. The world factbook: Afghanistan. https://www.cia.gov/library/publications/the-world-factbook/geos/af.html, 1 March 2011. Accessed 1 March 2011.
12. James Allen, Andrew D. James, and Phil Gamlen. Formal versus informal knowledge networks in r&d: a case study using social network analysis. *R&D Management*, 37(3):179–196, 2007.
13. Alon Altman and Moshe Tennenholtz. Ranking systems: the PageRank axioms. In *Proceedings of the 6th ACM conference on Electronic commerce*, EC '05, pages 1–8, New York, NY, USA, 2005. ACM.
14. Vladimir Batagelj and Andrej Mrvar. Pajek: Program for large network analysis. http://vlado.fmf.uni-lj.si/pub/networks/pajek/, 25 September 2010. accessed 30 January 2011.

15. Ramiro Berardo. Processing complexity in networks: a study of informal collaboration and its effect on organizational success. *The Policy Studies Journal*, 37(3):pp. 521–539, 2009.

16. Phillip Bonacich. Some unique properties of eigenvector centrality. *Social Networks*, 29:p. 555–564, 2007.

17. Stephen P. Borgatti. Structural holes: Unpacking burt's redundancy measures. *Connections*, 20(1):35–38, 1997.

18. Stephen P. Borgatti. Network measures of social capital. *Connections*, 21(2):27–36, 1998.

19. Ronald S. Burt. *Structural Holes: The Social Structure of Competition*. Harvard University Press, Cambridge, Massachusetts, 1992.

20. Ronald S. Burt. *Neighbor Networks*. Oxford University Press, New York, NY, 2010.

21. Kathleen M. Carley, Dave Columbus, Mike Bigrigg, Jana Diesner, and Frank Kunkel. *AutoMap User's Guide*. Institute for Software Research School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, June 2009.

22. James S. Coleman. *Foundations of Social Theory*. First Harvard University Press, 1994.

23. Independent Election Commission. Afghanistan 2010 Wolesi Jirga election final certified results. http://www.iec.org.af/results_10/eng/, 2010. Accessed 28 February 2011.

24. W. J. Conover. *Practical Nonparametric Statistics*. John Wiley and Sons, New York, NY, 2^{nd} edition, 1980.

25. Jock Covey, Michael J. Dziedzic, and Leonard R. Hawley. *The Quest for Viable Peace: International Intervention and Strategies for Conflict Transformation*. United States Institute of Peace, 1200 17th Street NW, Suite 200, Washington, DC 20036, 2^{nd} edition, 2006.

26. Maj Gen Scott S. Custer. CJCS notice 3245.01: Military support for stability, security, transition, and reconstruction SSTR operations, 12 May 2006. accessed 11 Nov 2010.

27. Department of Defense, Washington DC. *Military Support to Stabilization, Security, Transition, and Reconstruction Operations Joint Operating Concept*, 2^{nd} edition, December 2006.

28. Brunz V. Erreck. Afghan bios: Who is who in afghanistan. http://www.afghan-bios.info, March 2011. Accessed on 4 March 2011.

29. Michéle Flournoy. DoDD 3000.05 stability operations. http://www.dtic.mil/whs/directives/corres/pdf/300005p.pdf, 16 September 2009. accessed 15 November 2010.

30. Hans T.W. Frankort. Structural holes, technological resources, and innovation: A study of an interfirm r&d network. *Academy of Management Proceedings*, pages 1 – 6, 2008.

31. Ashraf Ghani and Clare Lockhart. In *Fixing Failed States: A Framework for Rebuilding our Fractured World*, Authors@Google, page 241, 76 Ninth Ave, 4^{th} Floor, New York, NY, June 2008.

32. David C. Gompert, Terrence K. Kelly, Brooke Stearns Lawson, Michelle Parker, and Kimberly Colloton. *Reconstruction under fire: Unifying Civil and military counterinsurgecy*. RAND Corporation, 1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407, 2009. www.rand.org.

33. Mark S. Granovetter. The strength of weak ties. *American Journal of Sociology*, 78(6):1360–1380, May 1973.

34. Casper Grathwohl, editor. *New Oxford American Dictionary*. Oxford University Press, 2^{nd} edition, 2005. http://oxforddictionaries.com/.

35. Aric Hagberg. Network x overview. http://networkx.lanl.gov/overview.html#who-uses-networkx, Aug 2010. accessed on 11 Jan 2011.

36. Aric Hagberg. Network x reference list. http://networkx.lanl.gov/reference/index.html, Aug 2010. accessed on 11 Jan 2011.

37. Robert A. Hanneman and Mark Riddle. *Introduction to Social Network Methods*. University of California, Riverside, CA, 2005. published in digital form at http://faculty.ucr.edu/~hanneman/.

38. Headquarters, Department of the Army, Wasington D.C. *Field Manual 3-24 Counterinsurgency*, 16 June 2006. http://www.army.mil/usapa/doctrine/active_fm.html.

39. Headquarters Department of the Army, Washington DC. *FM 3-07: Stability Operations*, October 2008.

40. Sonali Huria. Failed states and foreign military intervention: The afghanistan imbroglio. Special Report 67, Institute of Peace and Conflict Studies, B-7/3, Safdarjung Enclave, New Delhi 110029, February 2009.

41. Al Jazeera. Afghan poll results delayed. http://english.aljazeera.net/news/asia/2010/10/201010182281827619.html, 18 October 2010. Accessed 28 February 2011.

42. Joint Chiefs of Staff, Washington DC. *Joint Publication 3-24: Counterinsurgecy Operations*, 5 October 2009.

43. Joint Chiefs of Staff. *Joint Publication 3-0: Joint Operations*, 2^{nd} edition, 22 Mar 2010. Original publication date of 17 Sep 2006, http://www.dtic.mil/doctrine/new_pubs/jp3_0.pdf.

44. Seth G. Jones, Lee H. Hilborne, C. Ross Anthony, Lois M. Davis, Federico Girosi, Cheryl Benard, Rachel M. Swanger, Anita Datar Darten, and Anga Timilsina. *Securing Health: Lessons from nation-building missions*. RAND Corporation, 1200 South Hayes Street, Arlington, VA 22202, 2006.

45. Dr. Jack D. Kem. Stability and reconstruction operations: Connecting the dots between military and civilian efforts. *Small Wars Journal*, 7, Feb 2007.

46. Muhammad Wais Khitab. Many Karzai rivals find way into Parliament. http://www.pajhwok.com/en/2011/01/22/many-karzai-rivals-find-way-parliament, 22 January 2011. Accessed 28 February 2011.

47. David Knoke and Song Yang. *Social Network Analysis*. Number 154 in 07. SAGE Publications, 2^{nd} edition, 2008.

48. Clyde J. Mitchell. *Social networks in urban situations: Analysis of personal relationships in Central African Towns*. Manchester University Press, 1969.

49. Greg Mortenson and David Oliver Relin. *Three cups of tea: one man's mission to promote peace one school at a time*. Penguin Books, 80 Strand, London WC2R 0RL, England, 2006.

50. National Research Council. *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*. National Academies Press, 500 Fifth St. N.W. Washington D.C., 20001, 2008.

51. Andrew Natsios. Fragile states strategy. Technical report, U.S. Agency for International Development, 2005. http://www.usaid.gov/policy/2005_fragile_states_strategy.pdf.

52. BBC News. Afghanistan tribal map. http://news.bbc.co.uk/hi/english/static/in_depth/world/2001/war_on_terror/key_maps/ethnic_pashtun.stm, 28 Sep 2010.

53. Rod Nordland. Karzai agrees to seat new Afghan Parliament. http://www.nytimes.com/2011/01/23/world/asia/23afghan.html, 22 January 2011. Accessed 28 February 2011.

54. iGraph Project. The iGraph library. http://igraph.sourceforge.net/index.html, 2009. accessed 30 January 2011.

55. Abdullah Qazi. Biography: Zarar Ahmad Moqbel. http://www.afghan-web.com/bios/moqbel.html, 10 May 2010. Accessed 28 February 2011.

56. Hamid Shalizi. Karzai opens afghan parliament, taunts west. http://news.yahoo.com/s/nm/20110126/wl_nm/us_afghanistan_parliament, 26 January 2011. Accessed 28 January 2011.

57. Human Terrain System. My cousin's enemy is my friend: a study of pashtun "tribes" in afghanistan. Technical report, United States Army, TRADOC G2, Fort Leavenworth, KS, September 2009. http://easterncampaign.files.wordpress.com/2009/11/my-cousins-enemy-is-my-friend-a-study-of-pashtun-tribes.pdf.

58. U.S. Department of State. *Afghanistan: a country study*. Area handbook series. Claitor's Publishing Division, 3165 S. Acadian at I-10, Baton Rouge, LA, 2001.

59. U.S. Department of State. Afghan national assembly. http://merln.ndu.edu/archivepdf/afghanistan/State/60088.pdf, January 2006. Accessed 18 December 2010.

60. Renée van der Hulst. Introduction to social network analysis (sna) as an investigative tool. *Trends in Organized Crime*, 12:101–121, 2009. 10.1007/s12117-008-9057-6.

61. Diedrik van Liere, Jasper Voskuilen, and Aric Hagberg. Structural holes in python. https://networkx.lanl.gov/trac/ticket/191, September 2010.

62. Guido van Rossum. Python patterns - implementing graphs. http://www.python.org/doc/essays/graphs/, 1998.

63. Dennis D. Wackerly, William Mendenhall III, and Richard L. Scheaffer. *Mathematical Statistics with Applications*. Thomson West, 10 Davis Drive, Belmont, CA, 94002, 7^{th} edition, 2008.

64. Gordon Walker, Bruce Kogut, and Weijian Shan. Social capital, structural holes, and the formation of an industry network. *Organization Science*, 8(2):109–125, March-April 1997.

65. Stanley Wasserman and Katherine Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge, 1998. ID: 441070386.

66. Noah Webster. *Merriam-Webster's Collegiate Dictionary*. Encyclopedia Britannica, Springfield, Massachusetts, 10^{th} edition, 1999.

67. John Western, Robert Stimson, Scott Baum, and Yolanda Van Gellecum. Measuring community strength and social capital. *Regional Studies*, 39(8):1095–1109, November 2005.

68. Michael Woolcock. Social capital and economic development: Toward a theoretical synthesis and policy framework. *Theory and Society*, 27:151–208, 1998.

69. Xinhua. List of Afghan parliamentary candidates finalized. http://english.people.com.cn/90001/90777/90851/7035773.html, 23 June 2010. Accessed 28 February 2011.

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 074-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
24-03-2011	Master's Thesis	August 2010 - March 2011

4. TITLE AND SUBTITLE Utilizing Social Network Analysis in Support of Nation Building	**5a. CONTRACT NUMBER**
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Brandon J. Bernardoni, Capt., USAF	**5d. PROJECT NUMBER**
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 Hobson Street, Building 642 WPAFB OH 45433-7765	**8. PERFORMING ORGANIZATION REPORT NUMBER** AFIT/OR/MS/ENS/11-01
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Mr. August G. Jammarous USSOCOM HQ 7701 Tampa Point Blvd Bldg. 501 A, Room 121, 351 MacDill AFB, FL 33621	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Social network analysis is a powerful set of techniques used by social scientists to study the formal and informal interrelations in a community. Since 9/11 these techniques have been increasingly utilized by the defense and intelligence communities to analyze terrorist networks to aid in thwarting our foes. This study investigates the use of social networks and structural holes to facilitate nation building in failed and failing states. Investigation of the underlying social structure of a community, identifying structural holes and gaps within the government, Security Stabilization Transition and Reconstruction Operations (SSTRO) efforts can be focused to strengthen the host nation government to provide security and unity for its citizens.

This investigation focused on exploring techniques that link individuals in the professional and governmental community. It was found that Burt's technique of structural holes can be applied to a national level in order to identify structural gaps within an ethnically fractured failing state. This technique can highlight national, regional, or local holes that can be filled to facilitate nation building.

15. SUBJECT TERMS

Social Network Analysis, Structural Holes, Nation Building, Failed and Failing States, Stability Security Transition and Reconstruction Operations (SSTRO), Afghanistan.

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON** Dr. Richard Deckro, AFIT/ENS
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(Include area code)* 937-255-3636 x4325
U	U	U	UU	170	